Summer 2019

Targeting the Pin Domain of Type II Toxins Is a Novel Approach to Treat Infections Caused by Nontypeable *Haemophilus Influenzae*

Ashley Lynne Molinaro
Old Dominion University, asanf003@odu.edu

TARGETING THE PIN DOMAIN OF TYPE II TOXINS IS A NOVEL APPROACH TO TREAT INFECTIONS CAUSED BY NONTYPEABLE *HAEMOPHILUS INFLUENZAE*

by

Ashley Lynne Molinaro

B.S. May 2010, Norwich University

A Dissertation Submitted to the Faculty of
Old Dominion University in Partial Fulfillment of the
Requirements of the Degree of

DOCTOR OF PHILOSOPHY

BIOMEDICAL SCIENCES

OLD DOMINION UNIVERSITY

AUGUST 2019

Approved by:

Dayle A. Daines (Director)

Fred C. Dobbs (Member)

Christopher J. Osgood (Member)

Jing He (Member)

ABSTRACT

TARGETING THE PIN DOMAIN OF TYPE II TOXINS IS A NOVEL APPROACH TO TREAT INFECTIONS CAUSED BY NONTYPEABLE *HAEMOPHILUS INFLUENZAE*

Ashley Lynne Molinaro
Old Dominion University, 2019
Director: Dr. Dayle A. Daines

Toxin-antitoxin (TA) gene pairs have been identified in nearly all bacterial genomes sequenced to date and are thought to facilitate persistence and antibiotic tolerance. TA loci are classified into various types based upon the characteristics of their antitoxins, with those in type II expressing proteic antitoxins. Many toxins from type II modules are ribonucleases that maintain a PilT N-terminus (PIN) domain containing conserved amino acids considered essential for activity. The *vapBC* (*v*irulence *a*ssociated *p*rotein) TA system is the largest subfamily in this class and has been linked to pathogenesis of nontypeable *Haemophilus influenzae* (NTHi). This dissertation presents three studies investigating type II TA modules in NTHi. The first study determined the crystal structure of the VapBC-1 complex. Based on this structure, aspartate-to-asparagine and glutamate-to-glutamine mutations of four conserved residues in the PIN domain of VapC-1 were constructed, and the effects of these mutations on pathogenesis were tested *ex vivo*. A novel model system was designed that consisted of an NTHi Δ*vapBC-1* strain complemented *in cis* with the toxin mutants in tandem with the wild-type antitoxin controlled by the *vapBC*-1 native promoter in single copy. This determined that a single mutation to a conserved amino acid in the PIN domain significantly reduced the survival of NTHi. The second study measured the induction of a type II antitoxin *in trans* and the induction of a type II toxin *in cis* with or

without an *E. coli* arabinose permease *in cis* in the background of NTHi. A bioassay was then designed to investigate the ability of NTHi to transport arabinose, the molecule necessary to induce the *araBAD* genes. Lastly, a bioinformatics study was performed on 12 recently-sequenced clinical isolates of NTHi to determine the conservation of four *vap* operons. Seven strains maintained all four operons, the other five strains maintained three and the operons all shared significant sequence homology. These three studies provide insight into the importance of the type II TA loci in NTHi and due to their conservation and contribution to pathogenesis, these modules prove to be promising candidates for the development of antimicrobial treatments.

This dissertation is dedicated to my daughters, Paige and Claire. Having you during this journey has grounded me and brought me more strength than I could have ever imaged. I urge you to follow your dreams and reach higher than you ever thought possible. I love you more than you will ever know.

ACKNOWLEDGMENTS

I could not have completed this dissertation without the assistance of many people. I would like to first thank my advisor Dr. Dayle Daines, for accepting me into her lab five years ago, a new Lieutenant fresh from a deployment, looking to pursue a Master's Degree. Little did I know my passion for research would grow and I would finish this journey with a Doctorate. Thank you to my committee members: Drs. Fred Dobbs, Jing He, and Christopher Osgood for their guidance and thoughtful advice.

I would like to acknowledge Harmony Hancock-Martel, my best friend and science sister. I'm not quite sure why the stars aligned and brought us together during such different parts of this program but I will forever be grateful. Thank you for the many hours of complaining, encouragement, advice, tears, and laughter. I am appreciative to the dozens of people who watched my girls on days they were home sick from daycare and I could not leave work or the many weekends I spent in the lab. It really takes a village. Thank you to my parents, Deb and Jeff Sanford, for being my very first cheerleaders through life.

Finally, I would like to recognize my husband, Rob. Two babies, two deployments, multiple moves, and this grueling program, have challenged us these last five years in ways that would make any couple split. Thank you for your ever-loving support and sometimes tough love, without you I would not have completed this program.

This research was funded by grants to Dr. Dayle Daines from the National Institute on Deafness and other Communication Disorders R01 DC010187 and cooperative agreement U01 DC014756 to DAD.

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

Page

OVERVIEW

Haemophilus influenzae was originally called *Bacillus influenzae* (or Pfeiffer's bacillus) by bacteriologist Richard Pfeiffer when it was mistakenly identified as the causative agent of the 1889 influenza pandemic (1). It was not until the 1920s that the organism was named *H. influenzae* (2). It is a Gram-negative obligate commensal of the upper respiratory tract. *H. influenzae* is divided into two groups – encapsulated strains and un-encapsulated or nontypeable strains (NTHi). There are six capsular serotypes, a-f (3), whose capsular region (*cap*) all share the same organization (4-7). The *cap* locus is comprised of three regions – region I contains the *bexDCBA* genes which code for an ATP-driven polysaccharide export apparatus (8), region II contains serotype-specific DNA which distinguishes the six serotypes from each other (9), and region III contains the *hcsAB* genes which facilitate transport of the capsular polysaccharide across the outer membrane (Fig. 1) (10). *H. influenzae* stereotype b (Hib) strains

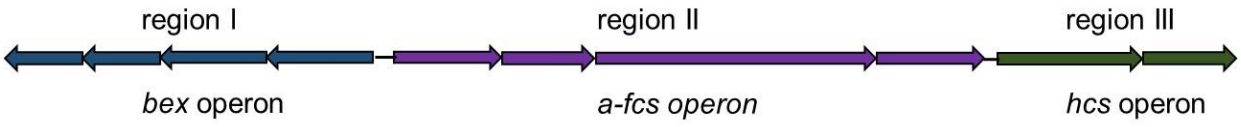

FIG 1 Representation of the *cap* locus in *H. influenzae*. The polysaccharide capsule in typeable *H. influenzae* is controlled by the *cap* locus for all six serotypes (a-f). Regions I and III are conserved among serotypes while region II encodes the genes responsible for the specific polysaccharide synthesis. Region II can vary in the number of genes responsible for the biosynthesis of each serotype. This figure is adapted from reference 195.

were most often associated with invasive diseases such as meningitis, cellulitis, and septicemia, related to a high incidence of morbidity and mortality (11). However, since the approval of the first Hib conjugate vaccine in 1987, there has been a significant decrease in disease associated with Hib (12, 13).

H. influenzae is a human-adapted commensal of the upper respiratory tract. Meningitis caused by Hib has significantly decreased in countries that routinely vaccinate against the bacterium. However, in developing countries it is estimated that meningitis affects 7-13 million children under the age of five annually (14). NTHi is the most common bacterial cause of chronic obstructive pulmonary disease (COPD), which according to the World Health Organization (WHO), is the third leading cause of death worldwide (15, 16). NTHi is also a leading cause of recurrent otitis media (OM) in the pediatric population (17). The financial implications of these three diseases is in the billions of dollars annually (18-20).

Unfortunately, the Hib vaccine is not effective in preventing infections caused by NTHi because the conjugate vaccine links a carrier protein to the polyribosylribitol phosphate (PRP) capsule unique to Hib. Due to the absence of a capsule as well as the genetic diversity of NTHi (21, 22), a suitable vaccine candidate against infections caused by NTHi is not available. Consequently, NTHi infections have been on the rise since the implementation of the Hib vaccine. NTHi strains share a number of virulence factors such as lipooligosaccharide (LOS), IgA protease, biofilm formation, adhesions, and other surface structures; however, the heterogeneity of the species makes it increasingly difficult to develop a vaccine suitable against a large number of isolates (23). Some of the outer membrane proteins that have been explored as potential

vaccine candidates include: P1, P2, P4, P6, Tbp1, Tbp2, and protein D (24). Two vaccines seemed to provide positive results: a ketodeoxyoctanoate (KDO) based glycoconjugate vaccine elicits a response from the monoclonal antibody 6E4 and binds to KDO (25). However, susceptible strains would have to express the KDO epitope in high levels, and in a small study of 33 NTHi strains, only 36% of the isolates were killed by the anti-KDO antibody (25). Another potential candidate is the protein D conjugate vaccine. In a clinical trial of nearly 5,000 infants, protective efficacy was reported in 36% of infants vaccinated with the 10-valent pneumococcal *H. influenzae* protein D conjugate vaccine (PHiD-CV) (26). However, a number of studies measured carriage rates after children were fully vaccinated with the PHiD-CV and found no change in the rates (27-29).

While a vaccine would be ideal in the race to prevent diseases caused by NTHi, the antigenic variations among gene products in this species may just be too great for preparation of an effective vaccine. However, exploiting toxin-antitoxin (TA) modules is a novel way to treat disease and infections caused by NTHi. TA loci were first identified on plasmids as modules that aided in plasmid maintenance and stability by mediating the killing of plasmid-free cells and thereby ensuring stable plasmid inheritance (30, 31). Following the advent of whole genome sequencing, TAs were found to be maintained in the chromosomes of numerous Gram-negative and Gram-positive bacterial genera as well as in the Archaea (32-34). Chromosomally-located TAs have been implicated as important metabolic regulators in many human pathogens, including *Escherichia coli,* nontypeable *Haemophilus influenzae* (NTHi)*, Mycobacterium tuberculosis,* and *Staphylococcus aureus* (35-38). Stressful environmental conditions, such as nutrient

limitation, heat, pH, and oxidative stress from the host immune response or antibiotic therapy (39-42), initiate degradation of the labile antitoxin via bacterial proteases such as Lon and ClpXP (43, 44). After the antitoxin is degraded, the toxin is free to inhibit essential cellular processes, such as ATP and protein synthesis, DNA replication, peptidoglycan synthesis, and cell division (45-49), resulting in growth arrest and facilitating a persister state.

TA modules are operons typically organized with the antitoxin located directly upstream of the toxin. Six different TA families have been identified and characterized by the nature of the antitoxin. In type I and type III systems, the antitoxin is an RNA. The type I antitoxin is antisense RNA that binds the toxin mRNA under normal cellular conditions, inhibiting its translation (50). Following degradation of the type I antitoxin, the toxin is translated and depolarizes the cell membrane, thereby inhibiting ATP synthesis (50, 51). In the type III TA loci, the antitoxin is a pseudoknot RNA that binds to the protein toxin and prevents it from cleaving RNA (52, 53). Type II, IV, V, and VI systems all contain protein antitoxins. In type II systems, both the antitoxin and toxin are proteins that form a non-toxic complex upon translation (54). Many toxins of the type II TA modules have been characterized as ribonucleases (46, 55). Type IV systems are unique in that the antitoxin inhibits toxin activity by interacting directly with the targets of the toxin, which are cytoskeletal proteins necessary for proper cell division (49, 56). The only type V TA locus described to date is *ghoST*, in which the antitoxin inhibits the activity of the toxin by cleaving its mRNA (57). There is one identified type VI locus currently, designated as *socAB*. In this module, the antitoxin SocA does not inhibit the activity of the SocB toxin directly, but rather guides it to the protease ClpXP (58). When

SocB is not degraded by ClpXP, it binds to DnaN and inhibits DNA replication and elongation (58). Free-living prokaryotes tend to have more TA loci than obligate intracellular microorganisms, and pathogenic bacteria often have more TA loci than their non-pathogenic relatives (59). Maintaining numerous TA loci in their genomes may increase the ability of microorganisms to rapidly respond to stresses in their microenvironments, and there is growing interest in studying the contributions of these systems to microbial pathogenesis and persistence.

Type II TA loci contribute significantly to the survival and virulence of NTHi during infection (36, 60). With this in mind, this dissertation highlights the use of the newly solved crystal structure of VapBC-1 from NTHi to study the highly conserved PilT N-terminal (PIN) (61) domain and the implications it has on NTHi survivability in a recently published human infection model. A novel assay was designed in the background of NTHi to study the growth arrest and growth recovery of the type II toxin and antitoxin proteins utilizing IPTG and arabinose-inducible promoters. Lastly, 12 clinical isolates of NTHi were sequenced and the sequences were used to investigate the conservation and organization of four conserved type II TA loci maintained in *H. influenzae* strains. The conservation of these TA loci in NTHi makes them novel candidates for antimicrobial therapy.

CRYSTAL STRUCTURE OF VAPBC-1 FROM NONTYPEABLE *HAEMOPHILUS INFLUENZAE* AND THE EFFECT OF PIN DOMAIN MUTATIONS ON SURVIVAL DURING INFECTION

This chapter is published, and is reprinted here with permission:
Molinaro AL, Kashipathy MM, Lovell S, Battaile KP, Coussens NP, Shen M, Daines DA. 2019. Crystal Structure of VapBC-1 from Nontypeable *Haemophilus influenzae* and the Effect of PIN Domain Mutations on Survival during Infection. J Bacteriol 201.

INTRODUCTION

Type II modules are currently the best characterized TAs and the *vapBC* (*v*irulence *a*ssociated *p*rotein) system is the largest subfamily in this class, with VapB being the proteic antitoxin and VapC the ribonuclease toxin. The VapB antitoxin has two roles in the VapBC system: a region near its C-terminus binds to and neutralizes the toxin, while its N-terminus binds DNA and increases the TA complex stability in auto-regulating transcription of its cognate operon (Fig. 2) (62). The VapC toxin contains a conserved PIN domain that is crucial for its activity (61). While it has been shown that some VapC homologues are able to target specific RNA in bacterial cells (55, 63, 64), it is important to note that a VapC toxin has not yet been co-crystallized with its specific RNA target (62).

NTHi maintains two *vapBC* modules in its chromosome, designated as *vapBC-1* and *vapBC-2*. It has been previously shown that the *vapBC-1* locus significantly increases survival during infection, both *in vivo* and *ex vivo* (36). Despite low primary sequence homology among the VapC toxins from different microorganisms, the PIN domain is ubiquitous and characterized by highly conserved acidic, polar, and hydrophobic residues in the active site (62, 65). The VapBC-1 crystal structure was solved with the assistance of collaborators from the Protein Structure Laboratory at the University of Kansas and the

Advanced Photon Source at Argonne National Laboratory. Based on the VapBC-1 crystal structure, directed mutations were introduced to four PIN domain residues and evaluated the resulting mutant's protein-protein interactions, effects on the growth of *E. coli,* and ability to survive during infections of primary human tissues *ex vivo*. The results show that some of the mutations in the PIN domain of the VapC-1 toxin were associated with decreased toxicity in *E. coli*, but the mutants retained the ability to homodimerize and to heterodimerize with the wild-type cognate antitoxin, VapB-1. A new system was designed and constructed to quantify the effects of these mutations on NTHi survival during infections of primary human tissues *ex vivo*. Any mutation to a conserved amino acid in the PIN domain significantly decreased the number of survivors compared to that of the *in cis* wild-type toxin under the same conditions.

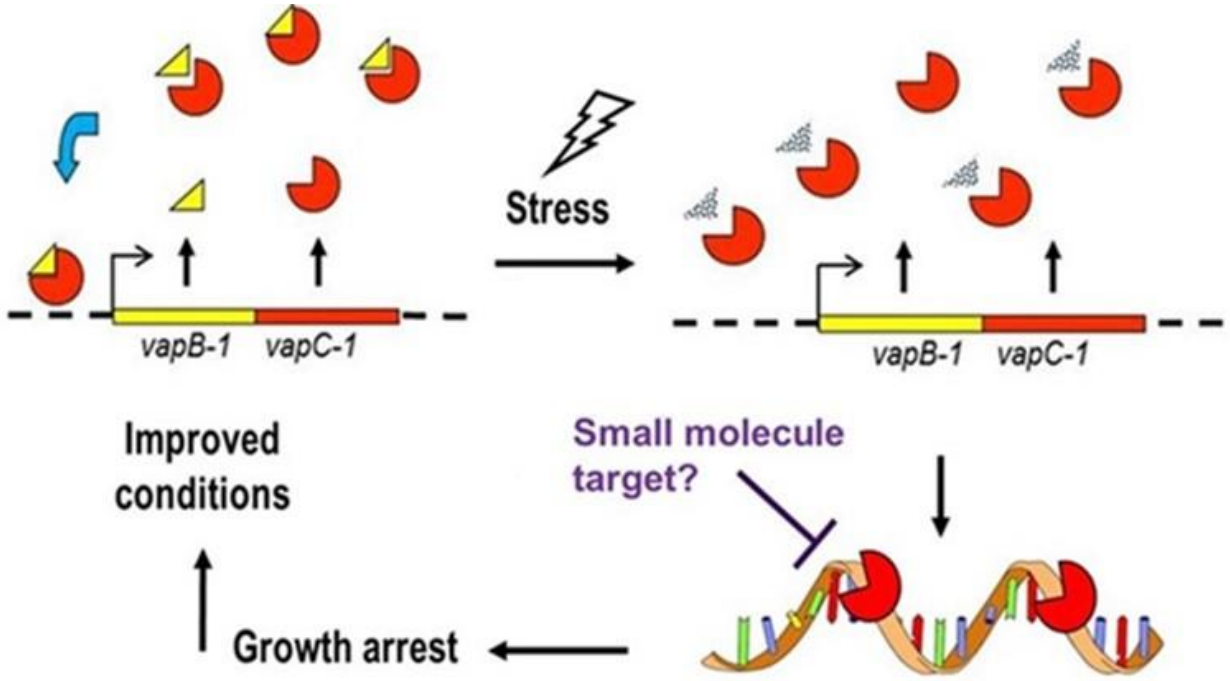

FIG 2 Model of TA locus function and inhibition of the toxin with a small molecule. Under stressful conditions, the antitoxin is degraded, freeing the toxin to cleave mRNA, resulting in bacterial growth arrest. Small molecules that inhibit the activity of the toxin VapC-1 would prevent growth arrest, thereby facilitating the complete clearance of a bacterial infection when used in conjunction with antibiotic therapy. This figure is adapted from reference 84.

RESULTS

Structural analysis of VapBC-1. Although VapBC structures from other organisms have been determined (66), the structure of VapBC-1 from NTHi has not been reported. Therefore, the crystal structure of VapBC-1 was determined from a polyhistidine-tagged construct that allowed the VapB-1 antitoxin to be co-expressed with the VapC-1 toxin. The primary sequences of the purified toxin and antitoxin used for crystallization are shown in Fig. 3.

A.

VapB-1

```
         10         20         30         40         50         60
MASMTGGQQM GRDPNSSSML TKVFQSGNSQ AVRIPMDFRF DVDTVEIFRK ENGDVVLRPV

         70         80         90
SKKTDDFLAL FEGFDETFIQ ALEARDDLPP QERENL
```

B.

VapC-1

```
         10         20         30         40         50         60
MIYMLDTNII IYLMKNRPKI IAERVSQLLP NDRLVMSFIT YAELIKGAFG SQNYEQSIRA

         70         80         90        100        110        120
IELLTERVNV LYPNEQICLH YGKWANTLKK QGRPIGNNDL WIACHALSLN AVLITHNVKE

        130        140
FQRITDLQWQ DWTKLEHHHH HH
```

FIG 3 Primary sequence of the VapBC-1 complex used for crystallization screening. (A) VapB-1 and (B) VapC-1. The underlined residues are from the pET24b cloning vector.

The overall structure of NTHi VapC-1 adopts an α-β-α sandwich fold in which five central β-strands are flanked by two α-helical regions typically observed for PIN-like domains (Fig. 4A). Subunits of VapC-1 form a dimer in which the hydrophobic surfaces of helices α3 and α5 are positioned approximately 10 Å from one another (Fig. 5A). The active site residues D6, E43, D99 and E120, depicted as spheres in Fig. 5A, are conserved amongst various VapC species and adopt a similar spatial arrangement relative to previously determined structures. The main network of hydrogen bonds is formed between dimer subunits through residues Y81 in helix α5 and N97, N98 and W101 of helix α6 (Fig. 5B). An additional hydrogen bond is present between E75 of subunit B and Y41 in helix α3 of subunit A. The reciprocal interaction, E75 in subunit A with Y41 of subunit B, is not observed as the side chains are positioned 5 Å apart. Thus, this hydrogen bond interaction is not likely a critical interaction between dimer subunits. Similar to previously determined structures (62, 67), the VapC-1 dimer has an interface area of 1,024 $Å^2$, a solvation free energy gain from dimer formation (ΔG_{int}) of -21.6 kcal/mol and 13% of the accessible surface area buried as determined with PISA (68). As noted above, there is minimal hydrogen bonding between the dimer subunits and no salt bridge interactions are present.

The VapC-1 dimer adopts a similar arrangement compared to structures from other organisms (Fig. 6). Comparison of the NTHi VapC-1 dimer with homologous structures, identified from a PDBeFold search (69) yielded RMS deviations between Cα-atoms of 1.85 Å (251 residues, VapC D98N from *Shigella flexneri*, PDB 5ECD) (70), 1.85 Å (253 residues, VapBC from *S. flexneri*, PDB 3TND) (71) and 1.79 Å (256 residues, VapBC

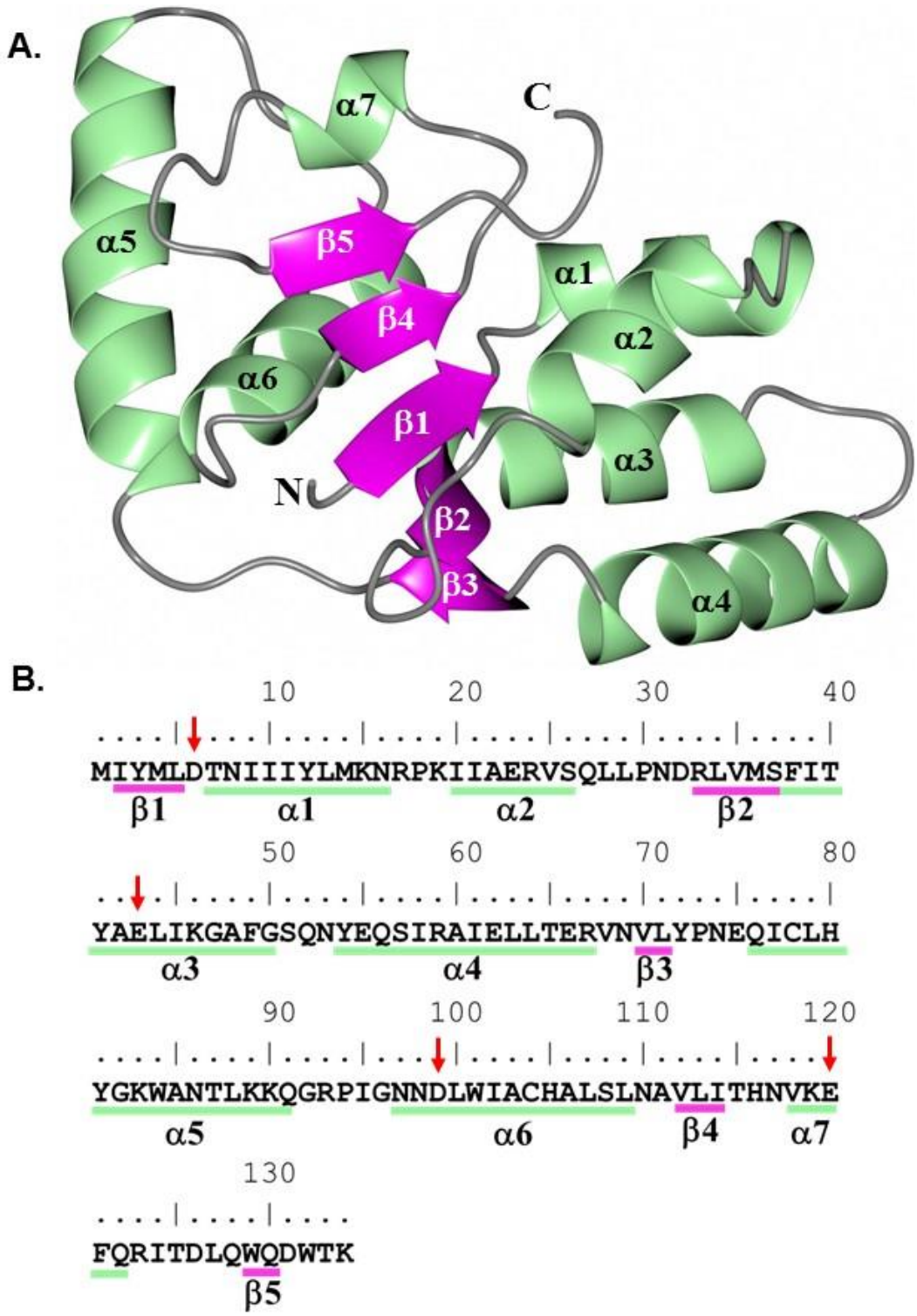

FIG 4 Structure of the *H. influenzae* VapC-1 monomer subunit. (A) Colored by secondary structure with α-helices (green) and β-sheets (magenta). (B) Primary and secondary structural elements of VapC-1. Conserved active site residues are indicated by the red arrows.

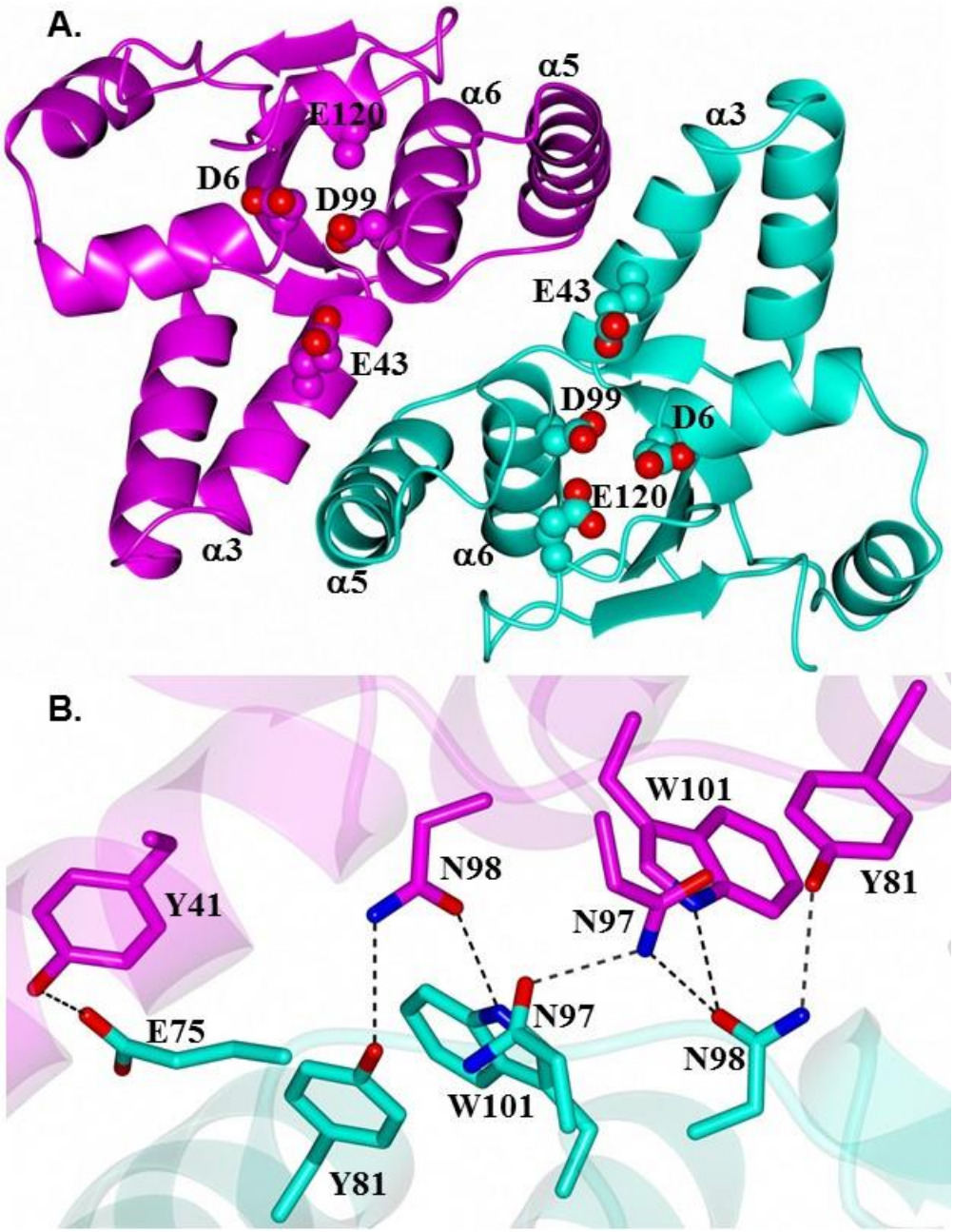

FIG 5 Structure of the *H. influenzae* VapC-1 dimer showing subunits A and B in magenta and cyan, respectively. (A) View along the 2-fold axis of the dimer. Helices that form the interacting surfaces between subunits are indicated. Conserved active site residues are rendered as spheres. Note that E120 (chain A) was partially disordered so all of the side chain atoms were not modeled. (B) Hydrogen bond interactions (dashed lines) between residues of the dimer.

from *Rickettsia felis* in complex with DNA, PDB 3ZVK) (72). The main difference is observed in helix α2 which is moved further away from helix α1 in the NTHi VapC-1 dimer as noted in panel A of Fig. 6. As shown below, this relative difference in the position of helices α1-α2 is necessary to accommodate VapB-1 binding. Slightly lower RMS deviations were observed when comparing single subunits and superposition with VapC-1 (subunit A) yielded the following: 5ECD (subunit A, 1.48 Å, 126 residues), 3TND (subunit C, 1.37 Å, 129 residues) and 3ZVK (subunit A, 1.54 Å, 131 residues). The main conformational differences occur in the relative positions of helices α1-α2 as was observed for superposition of the dimers noted above.

The VapB-1 portion of the NTHi VapBC-1 complex could be modeled from M1 to P72 (subunit C) and M1 to D69 (subunit D). The overall structure of VapB-1 consists of a ribbon-helix-helix motif in which the four β-sheets of the N-terminus dimerize with another subunit to form a β-barrel type architecture (Fig. 7). As determined using PISA (68), extensive hydrogen bonding and salt bridge interactions are present between the dimer subunits (27 and 11 respectively) which forms an interface area of 1,662 $Å^2$ and a solvation free energy gain of -17.0 kcal/mol. Notably, 25% of the accessible surface area is buried to form the dimer interface. The C-terminus contains a long stretch of flexible residues occupied by two α-helices that bind to the VapC-1 toxin forming a 1:1 interaction between VapB-1 and VapC-1 subunits. A similar architecture is observed for other VapBC structures, such as VapBC from *R. felis* in complex with DNA (PDB 3ZVK). The dimerization core regions are quite similar. However, the VapC-1 binding region of the C-terminus adopts a markedly different orientation relative to the VapB-1 β-barrel dimer core. This is not too surprising, given that VapBC complexes often adopt various

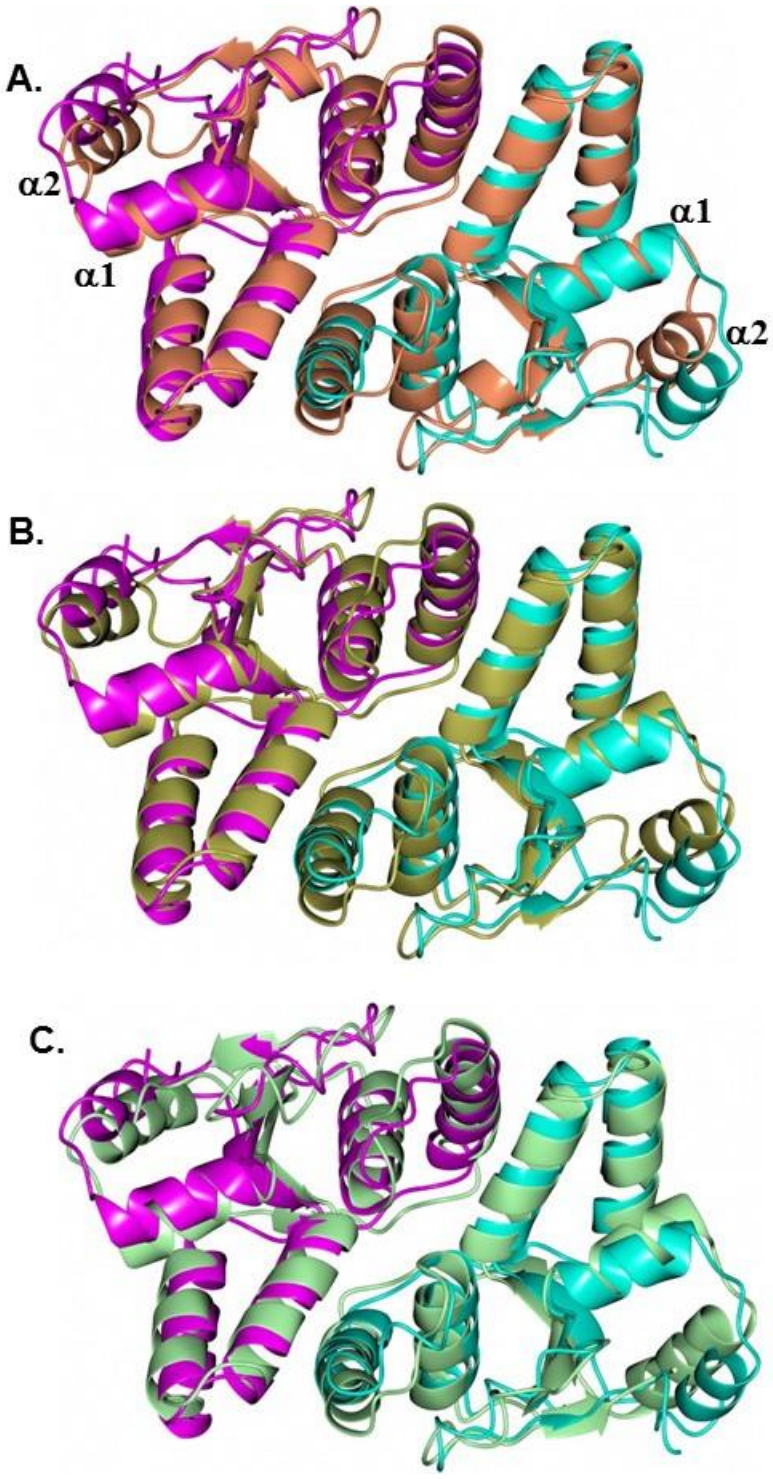

FIG 6 Comparison of the *H. influenzae* VapC-1 dimer (magenta/cyan) with the VapC dimer from homologous structures. (A) VapC D98N (coral) from *Shigella flexneri* (PDB 5ECD); (B) VapBC (gold) from *Shigella flexneri* (PDB 3TND); (C) VapBC (green) from *Rickettsia felis* in complex with DNA (DNA not shown) (PDB 3ZVK).

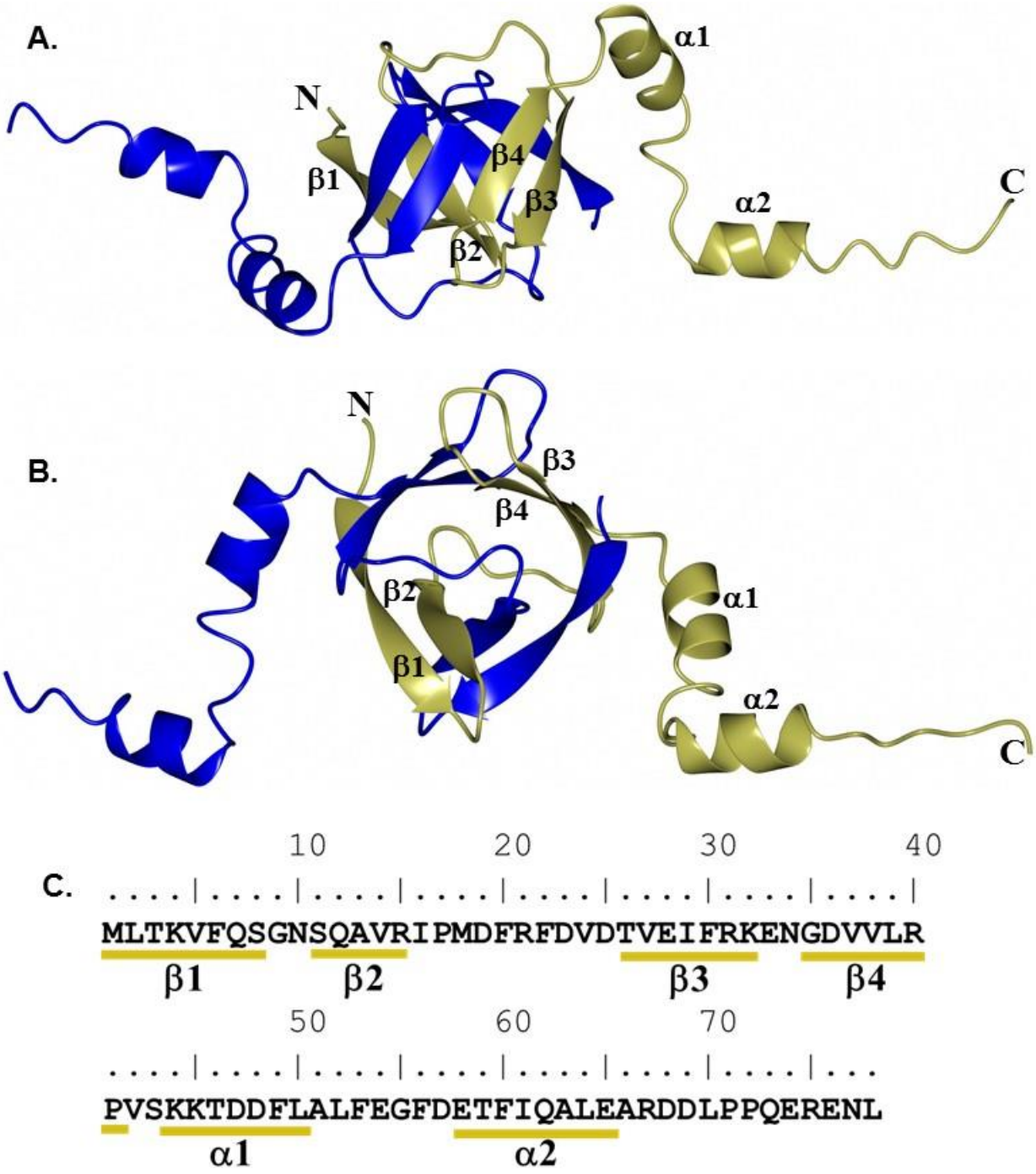

FIG 7 Structure of the *H. influenzae* VapB-1 dimer showing subunits C (gold) and D (blue). Secondary structure elements annotated for subunit C. (A) Perpendicular view; (B) view along the β-barrel axis; (C) primary and secondary structural elements of VapB-1.

quaternary structural arrangements that produce higher order oligomers (62).

Although the toxin:antitoxin form commonly observed consists of 1:1 interactions between subunits, the quaternary arrangement resulting from the binding of NTHi VapB-1 to VapC-1 in this structure appears to be unique. As shown in Fig. 8A, the asymmetric unit consists of a VapB-1 and VapC-1 dimer in which the C-terminal helices of VapB-1 (subunit C) occupy a cleft formed by α1-α4 of VapC-1 (subunit A). Subunit B of VapC-1 forms a similar interaction with VapB-1 (subunit D). However, this interaction is with a VapB-1 molecule related by crystallographic symmetry (Fig. 8B) and creates tightly packed pairs of VapB-1 and VapC-1 dimers (Fig. 8C). This arrangement results in a row of alternating VapB-1 and VapC-1 dimers.

As noted above, the C-terminal tails of VapB-1 are positioned within a cleft formed by α1-α4 of a VapC-1 dimer. Specifically, the C-terminal tail spanning P41-P72 (subunit C) and P41-D69 (subunit D) of VapB-1 (Fig. 9A) position the α-helices within a complementary cleft of VapC-1 subunits A and B respectively (Fig. 9B). The binding interface between VapB-1 and VapC-1 is 1,385 Å^2 with a solvation free energy gain of -21.8 kcal/mol. The majority of the hydrogen bond and salt bridge interactions occur at the C-terminus of VapB-1 spanning residues A66 to L70 (Fig. 10). Notably, R67 of VapB-1 interacts with E43 and D99 of VapC-1, which are conserved catalytic residues along with D6 and E120, amongst various VapC species (73). Similar interactions were observed for the *S. flexneri* VapBC structure (PDB 3TND), suggesting a similar mode of VapC-1 inhibition. Additional hydrogen bond interactions are observed between G55 and N52 of VapB-1 and N56 and N52 of VapC-1 (not shown).

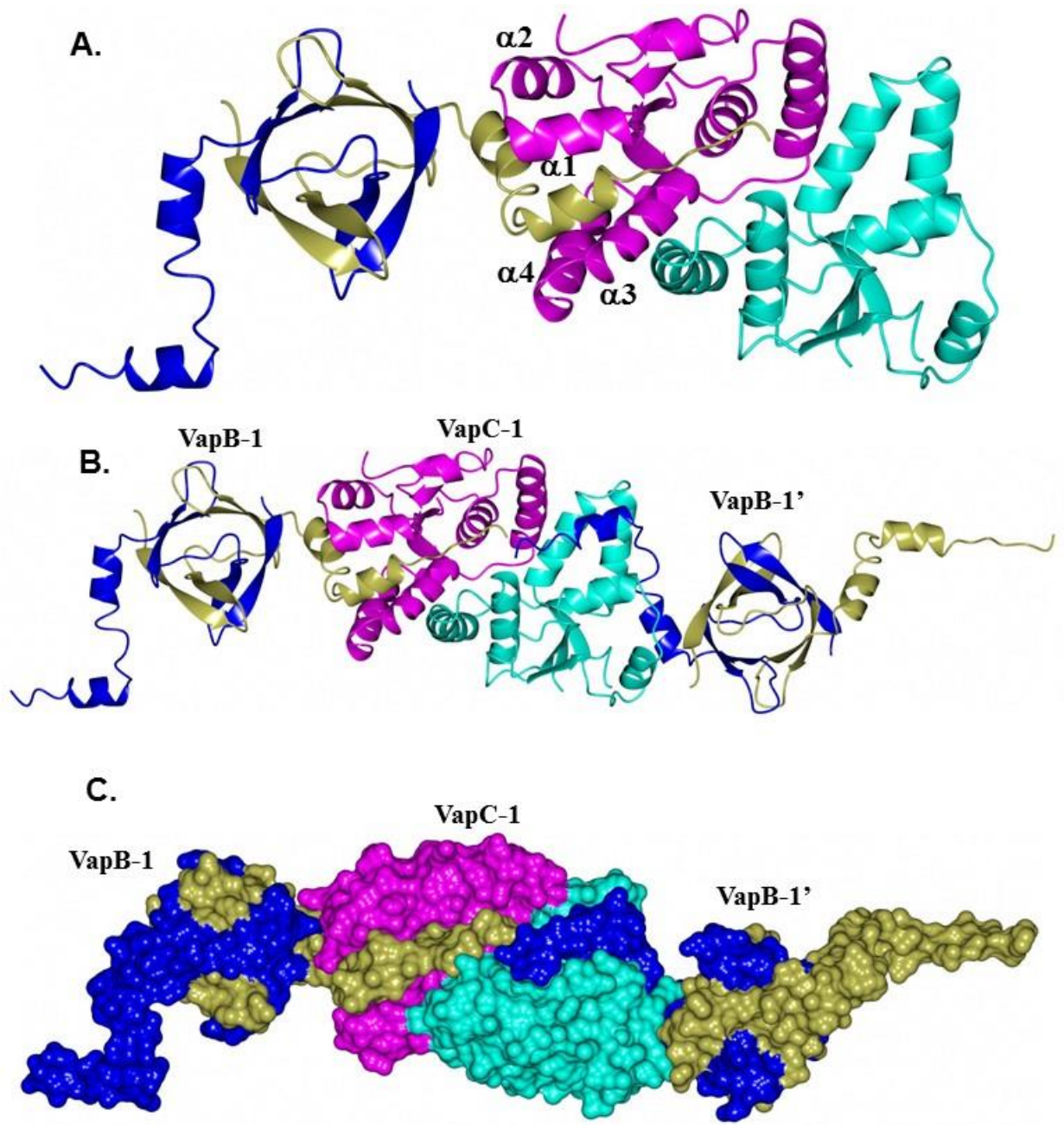

FIG 8 Structure of VapBC-1 with VapB-1 subunits C/D (gold/blue) and VapC-1 subunits A/B (magenta/cyan). (A) Asymmetric unit showing the positioning of the C-terminal tail of VapB-1 (subunit C) occupying the cleft formed by α1-α4 of VapC-1 (subunit A); (B) asymmetric unit with an extra VapB-1 dimer related by crystallographic symmetry (VapB-1′); (C) same as that shown in panel B but rendered as a surface.

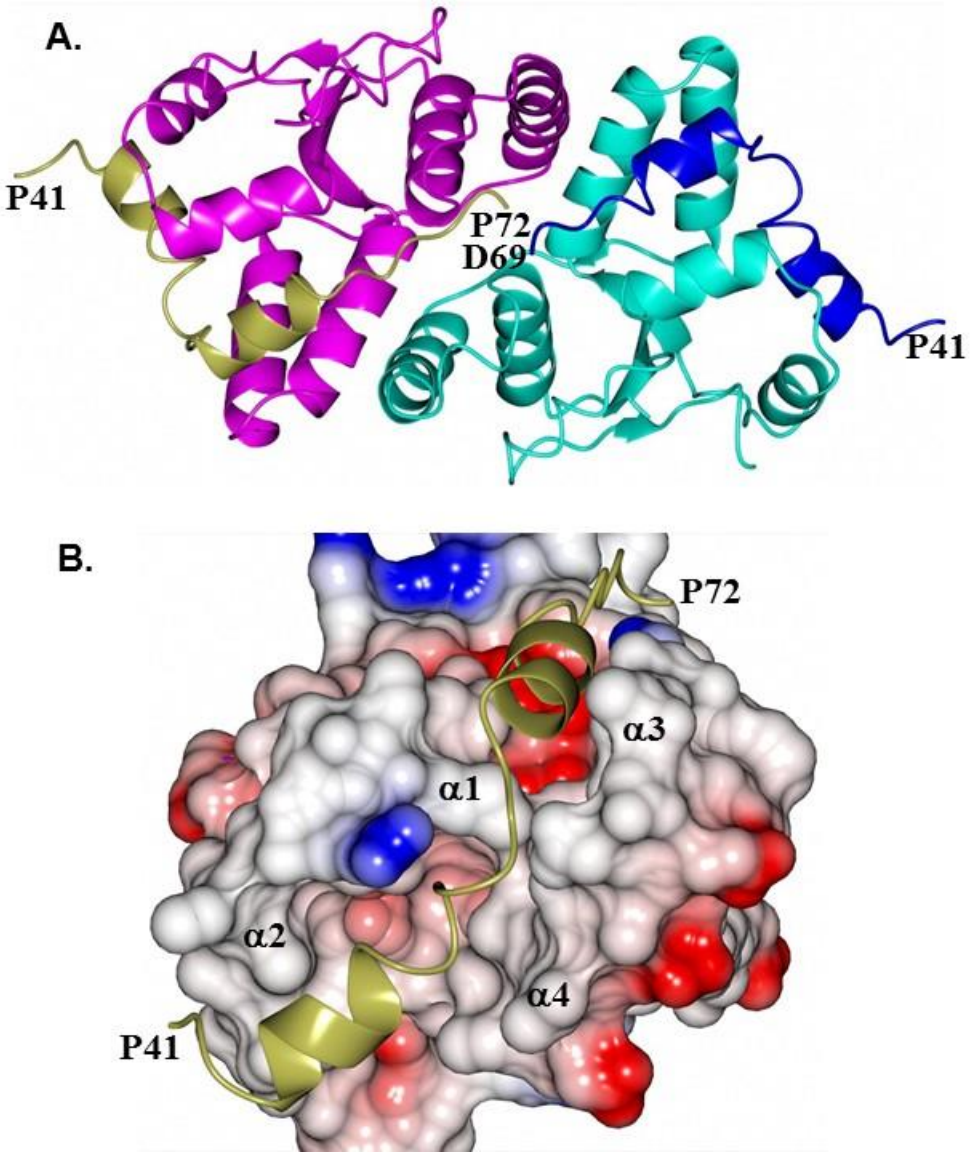

FIG 9 Binding of VapB-1 to VapC-1. (A) Zoomed-in view of Fig. 5B showing the binding of the C-terminal tail of VapB-1 subunit C residues P41to P72 (gold) and subunit D residues P41-D69 (blue) bound to the VapC-1 dimer; (B) electrostatic surface of VapC-1 subunit A showing the binding of the C-terminal α-helices of VapB-1 (P41-P72, gold) positioned within the cleft form by α1-α4 of VapC-1.

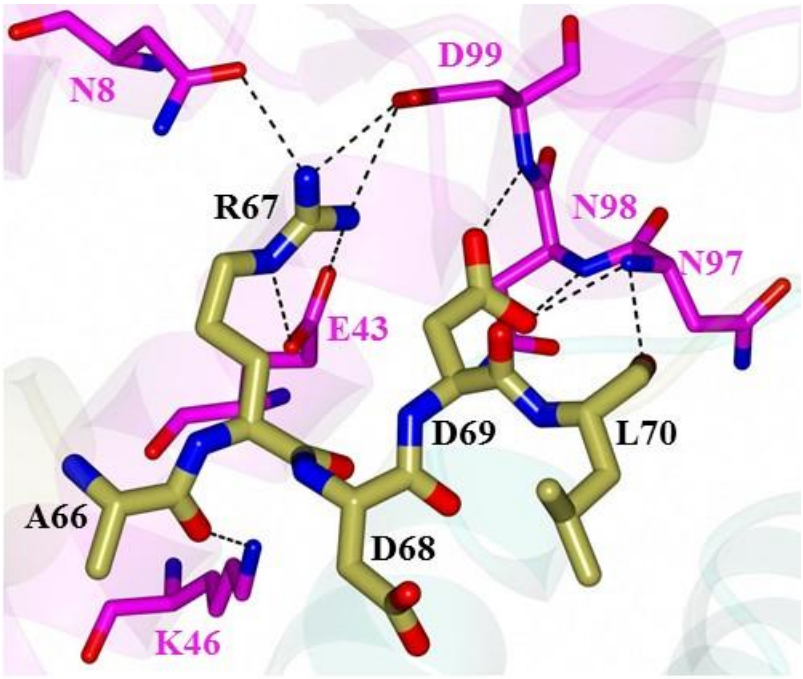

FIG 10 Hydrogen bond interactions between the C-terminal end of VapB-1 subunit C (gold) and VapC-1 subunit A (magenta).

Protein-protein interactions *in vivo.* Proteins encoded by type II TA modules form nontoxic heterocomplexes of the antitoxin and toxin homodimers under normal growth conditions *in vivo* (66). In order to examine antitoxin and toxin homodimer formation and heterodimerization, a bacterial LexA-based system (74) was employed. The LexA repressor consists of two domains, a DNA-binding domain (DBD) and a dimerization domain. The DBD can function as a transcriptional repressor only in dimeric form. Other domains can be fused in-frame to the DBD and will restore the repressor's function only if they interact (75). If the fused proteins interact, the *lacZ* reporter gene is repressed, reducing beta-galactosidase activity, which is quantitated biochemically as dimensionless Miller units (76). Strong protein-protein interactions result in minimal beta-galactosidase activity, whereas weak interactions result in significant activity.

This system was applied to evaluate the ability of VapC-1 single and double site-directed mutants to interact with the wild-type VapB-1 antitoxin (heterodimerization) as well as with themselves (homodimerization) *in vivo*. The homodimerization of the wild-type VapC-1 toxin was not evaluated due to its substantial toxicity when overexpressed from a high copy number plasmid. Table 1 shows that the VapC-1 mutants retained the ability to homodimerize with comparative strengths of interaction of E120Q > D6N > E43Q D99N > E43Q > D99N > D6N D99N. Mutant or wild-type VapC-1 toxins could heterodimerize with wild-type VapB-1, at relative levels of D6N > D6N D99N > E43Q D99N > E120Q > WT > E43Q > D99N (Table 2).

TABLE 1 Homodimerization of VapC-1 mutants

Construct	Homodimerization[a] of LexA DBD fusion in SU101 (Miller units)	
	Avg	**± SD**
pSR658 (control)	1302.6	78.2
VapC-1 D6N	38.5	12.9
VapC-1 D99N	59.5	9.4
VapC-1 D6N D99N	148.3	24.8
VapC-1 E43Q	55.5	9.8
VapC-1 E43Q D99N	39.2	2.2
VapC-1 E120Q	27.6	8.4

[a]The homodimerization of the wild-type VapC-1 toxin was not evaluated due to its substantial toxicity when overexpressed in high copy.

TABLE 2 Heterodimerization of VapC-1 mutants with the wild-type VapB-1 antitoxin

Construct	Heterodimerization of LexA DBD fusions in SU202 (Miller Units)	
	Avg	**± SD**
Control vectors	1813.4	130.0
VapC-1 D6N	37.9	11.5
VapC-1 D99N	360.5	30.6
VapC-1 D6N D99N	56.0	6.0
VapC-1 E43Q	202.2	18.9
VapC-1 E43Q D99N	113.3	5.5
VapC-1 E120Q	167.3	11.6
VapC-1 Wild-type	190.4	6.4

VapC-1 activity *in vivo.* A growth-recovery assay was used to evaluate VapC-1 activity in *E. coli*. Mutant or wild-type VapC-1 toxins were individually cloned under the control of the P_{BAD} promoter and induced with 3.3 mM (0.05% wt/vol) L-arabinose (77). The wild-type VapB-1 antitoxin was under the control of the P_{trc} promoter and its expression was induced with 0.5 mM IPTG (78). Thus, VapC-1 and VapB-1 could be induced either independently or simultaneously from compatible plasmids that are co-transformed and expressed in the same *E. coli* strain, LMG194. To initiate the assay, a culture in early growth phase was split and treated with either: (i) media only, (ii) arabinose-only, (iii) IPTG-only, or (iv) both arabinose and IPTG. Typically, significant growth arrest occurs at approximately 70 min after wild-type VapC-1 was induced, but this was not observed if the antitoxin is expressed alone (Fig. 11A, red and green symbols, respectively). If VapB-1 and VapC-1 were co-induced, the antitoxin inhibited the toxin ribonuclease activity and growth was similar to the non-induced control (Fig. 11A, purple and blue symbols, respectively). The results of experiments with the VapC-1 mutants are presented in Fig. 11B-G. Significantly attenuated growth was observed upon induction of VapC-1 in strains with the wild-type, D99N and E120Q mutant toxins (red symbols in Fig. 11A, C, and G, respectively), compared to the growth observed when VapB-1 was co-induced (purple diamonds). However, the VapC-1 D6N, D6N D99N, E43Q, and E43Q D99N mutants did not cause significant growth arrest when induced *in vivo* (Fig. 11B and D to F, red symbols).

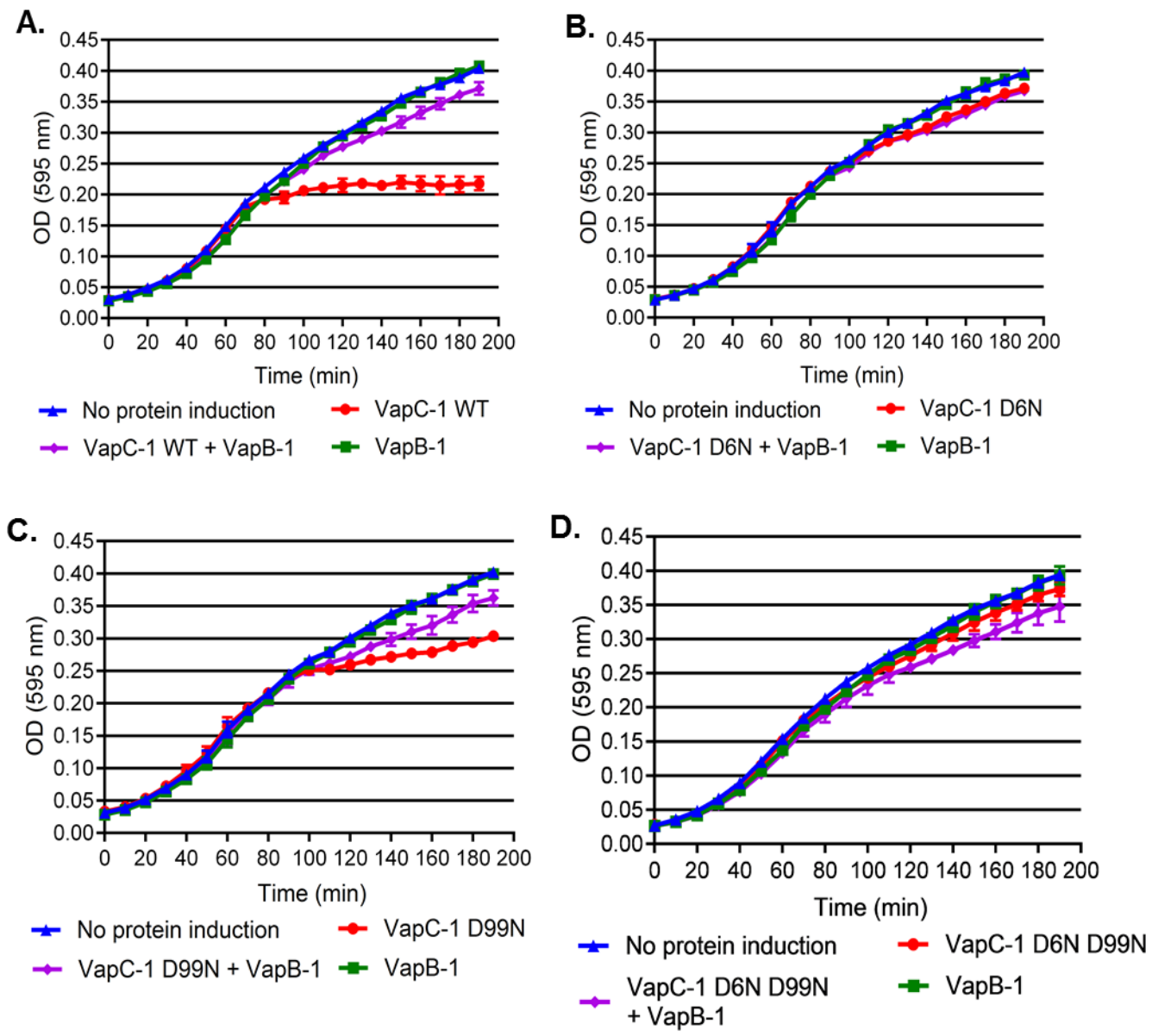

FIG 11 *E. coli* growth-recovery assays with inducible wild-type or mutant VapC-1 toxins. (A) Non-induced *E. coli* growth (measured by absorbance at 595 nm) or growth following induction of wild-type (WT) VapC-1, VapB-1, or both WT VapC-1 and VapB-1. (B to G) Similar growth experiments were performed with the VapC-1 D6N (B), D99N (C), D6N D99N (D), E43Q (E), E43Q D99N (F), and E120Q (G) mutants. All data are plotted as mean ± SD of the result of three biological replicates in triplicate. Statistical comparisons of growth following WT or mutant VapC-1 inductions against the corresponding non-induced controls were made using the repeated-measures ANOVA. The significance of the results were as follows: VapC-1 WT $P < 0.0001$ (A); D6N mutant $P = 0.21$ (B); D99N mutant $P = 0.0015$ (C); D6N D99N mutant $P = 0.15$ (D); E43Q mutant $P = 0.39$ (E); E43Q D99N mutant $P = 0.28$ (F); E120Q mutant $P = 0.017$ (G).

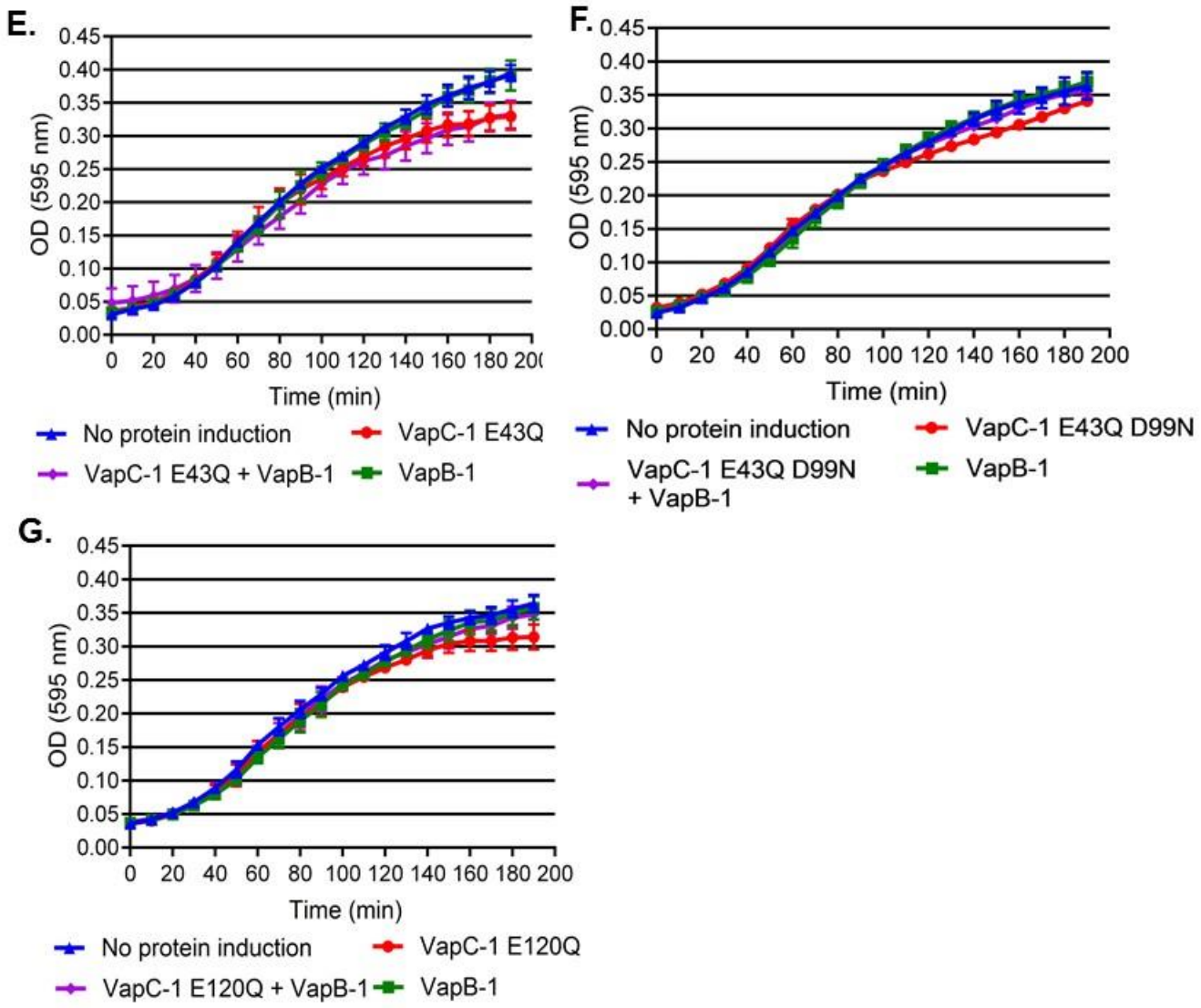

FIG 11 Continued.

Construction of an *in cis* ectopic system to study TA modules. The growth recovery assay in *E. coli* strain LMG194 is a valuable tool to evaluate the effects of toxin and antitoxin expression. However, the antitoxin and toxin proteins are cloned into separate multi-copy vectors to allow independent expression and are controlled by artificial promoters. Therefore, the assay does not allow either expression via the native transcriptional regulation of the locus or the characterization of a mutant locus in single copy on the chromosome. To evaluate these aspects, a novel system was developed that consisted of the native *vapBC-1* promoter controlling an operon containing the wild-type *vapB-1* gene followed by either the wild-type or each mutant *vapC-1* gene (Fig. 12). This system mimics the natural sequential arrangement and copy number of the operon. The module is flanked by homologous regions of the *metE* pseudogene from NTHi strain 86-028NP (NTHI_RS09270, an authentic frameshift resulting in a deletion of over 100 amino acids) which targets the reconstructed TA locus to a chromosomal region that is not in use. Briefly, fusions are amplified by PCR, used to transform NTHi, and are homologously recombined into the chromosome at the NTHI_RS09270 pseudogene site in single copy. NTHi is an obligate human parasite with natural deletions in many genes involving metabolism. Since it scavenges a number of essential molecules from its host, this results in the requirement to supplement growth media *in vitro* but not *in vivo* or *ex vivo* (79-81). The well-characterized 86-028NP Δ*vapBC-1* strain (36) was chosen as the background to analyze each tandem fusion to eliminate contributions by the wild-type *vapBC-1* locus. This novel approach is the most biologically relevant system currently available to study the effects of TA locus mutants in the background of NTHi.

To ensure that the *in cis* mutant modules had no effect on the replication of strains, the growth of the reconstructed wild-type, *vapBC-1* mutants, and the parent strain 86-028NP Δ*vapBC-1* was measured over 7 h. There were no significant differences in growth among the strains (data not shown), supporting the use of this model to investigate variations in pathogenesis attributable to mutations in VapC-1.

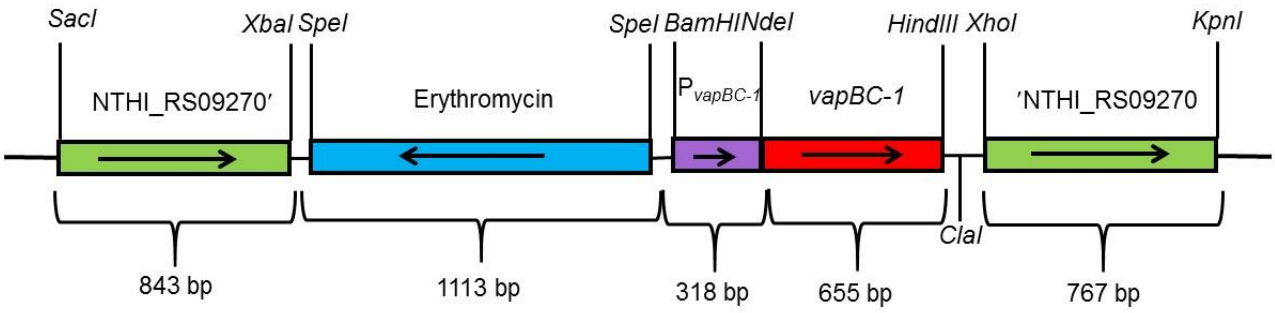

FIG 12 NTHi_RS09270 delivery vector for single-copy *in cis* expression of cloned genes. This construct targets the *metE* pseudogene in strain 86-028NP. The native *vapBC-1* promoter controls either the wild-type *vapBC-1* operon or an engineered operon consisting of the wild-type *vapB-1* antitoxin and one of the *vapC-1* mutated toxins.

Effect of VapC-1 mutations on survival of NTHi in primary human tissues. The EpiAirway model (AIR 100 ABF, MatTek, Ashland, MA, USA), which consists of antibiotic-free primary human respiratory epithelial tissue at the air-liquid interface, was used to evaluate the impact of VapC-1 mutations on NTHi survival during infections. After co-culturing the *in cis* complemented strains with EpiAirway tissues in antibiotic-free media for 48 h, the tissues were treated with 100 μg/ml gentamicin for 90 min to kill any

NTHi not internalized, as this antibiotic does not enter the tissues. The surviving internalized bacteria were recovered, diluted and plated for viable colony counts. This experiment was performed with n = 6 tissues originating from at least two different EpiAirway kits. The survival of all strains containing VapC-1 mutants were significantly reduced compared to the strain containing wild-type toxin (Fig. 13). The survival of the parent strain, 86-028NP Δ*vapBC-1*, was also significantly decreased, as previously reported (36).

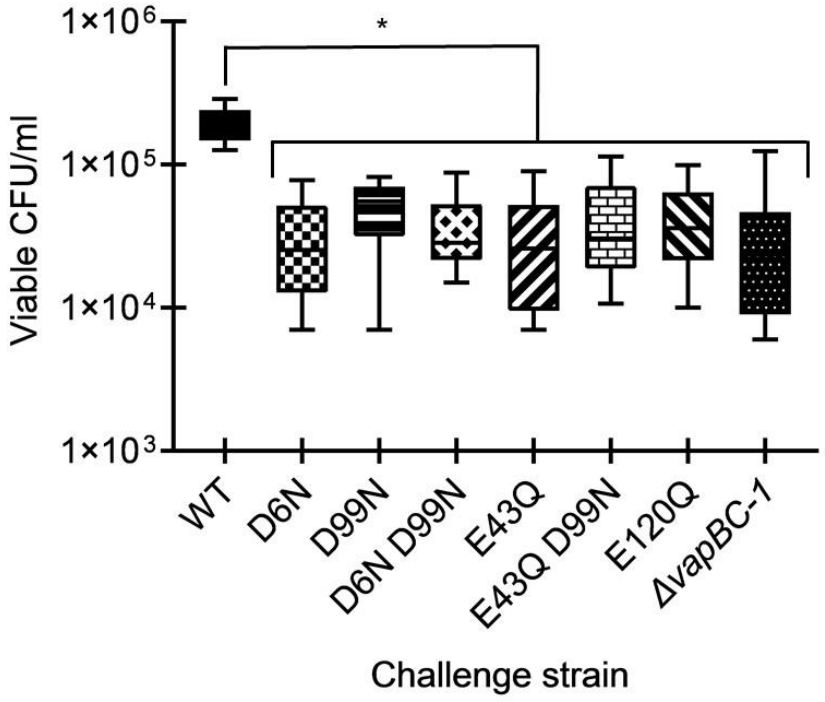

FIG 13 Number of viable CFU/milliliter of NTHi *in cis*-complemented mutants or the wild-type operon in primary human tissues at the air-liquid interface. EpiAirway tissues were infected at ~ 1.0 x 10^4 CFU per insert with either the 86-028NP *ΔvapBC-1* deletion strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (wild-type) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (D6N) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (D99N) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (D6N D99N) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (E43Q) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (E43Q D99N) strain, the 86-028NP *ΔvapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* (E120Q) strain, or the 86-028NP *ΔvapBC-1* parent strain. At 48 h after infection, gentamicin-resistant bacteria were harvested for colony counts. Data are expressed as the mean ± SD of six biological replicates in quadruplicate. All *in cis* mutants and the parent strain were significantly attenuated for survival compared to the strain carrying the wild-type VapC-1 *in cis*. An asterisk denotes a *P* of ≤ 0.05.

DISCUSSION

The VapC-1 toxin of NTHi is an efficient and highly active ribonuclease, and the *vapBC-1* locus is crucial for survival during infection, both *ex vivo* in primary human tissues, and *in vivo* in the chinchilla model of acute otitis media (36). The PIN domain of VapC homologues has been implicated in the binding and cleavage of RNA substrates (70) and many acidic residues in this domain are highly conserved (73). Other studies have evaluated various mutations to the PIN domain (82, 83), and it was previously demonstrated that the VapC-1 D99N mutant had significantly reduced ribonuclease activity, while retaining the ability to bind VapB-1 (84).

The NTHi VapBC-1 structure provides insight into heterodimeric interactions and enables comparisons with VapBC complexes from other organisms. Overall, the VapBC-1 homodimer subunits are similar to other reported VapB and VapC structures, adopting similar protein:protein interactions with the VapB C-terminal helices occupying clefts within VapC dimers. The homodimer subunits in VapBC complexes often form higher order oligomers such as heterotetramers, heterohexamers and heteroctamers (62). Interestingly, the VapBC-1 complex does not adopt any of these higher order quaternary structures but instead forms a unique linear array of alternating VapB and VapC dimers.

Based on the crystal structure of VapBC-1, two highly conserved aspartate and two glutamate residues in the toxin PIN domain were mutated to their beta-amino derivatives, asparagine and glutamine. Analyses of these mutant constructs indicated decreased ribonuclease activity *in vivo* measured by growth-recovery assays, and decreased survival during *ex vivo* infections of primary human tissues compared to the wild-type toxin when expressed in the background of NTHi using a novel single-copy *in*

cis system. This study also found that the effect of these mutations were not necessarily additive, as although a VapC-1 D99N mutant was significantly attenuated when overexpressed in *E. coli* growth-recovery assays, a double D6N D99N mutation was not. Interestingly, all VapC-1 mutants were able to strongly homodimerize as well as to bind wild-type VapB-1, suggesting that the module's quaternary structure is somewhat decoupled from the enzymatic activity of the toxin.

The results reported here support the notion that VapBC-1 (and TA modules in general) have evolved a number of mechanisms to maintain their activity and that their presence confers a significant advantage to the microorganism. Many studies have linked TA modules to the survival of bacterial pathogens in host tissues (36, 85-89). Achieving and maintaining persistence during an infection is central for both successfully evading the immune response and for nonspecific antibiotic tolerance, leading to a pathogen's survival of clinical therapy. This characteristic can substantially contribute to recurrent or chronic infections and implicates TA modules as therapeutic targets. Indeed, when the effects of the toxin mutations on NTHi survival inside primary human respiratory tissues *ex vivo* were tested, it was found that any perturbation of the acidic residues in the conserved PIN domain negatively affected survival compared to the wild-type toxin under the same conditions. These data suggest that targeting any of these individual residues in the toxin could potentially result in significantly decreasing the survival of NTHi during infection, and imply that small molecule inhibitors could be most effective in this approach.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The bacterial strains and plasmids used in these studies are listed in Table 3. *E. coli* strains were grown in LB broth or agar with or without 100 µg/mL ampicillin, 250 µg/ mL erythromycin, 20 µg/mL chloramphenicol, 30 µg/mL of kanamycin, or 12 µg/mL tetracycline, as required. NTHi strains were grown in brain heart infusion broth or agar supplemented with 10 µg/mL heme-histidine and 10 µg/mL β-NAD (sBHI), or on chocolate agar with 5 U/mL bacitracin. NTHi were routinely cultured at 37 °C with 5% CO_2. To construct the ectopic strains, transformants were selected on chocolate agar plates with 5 µg/mL erythromycin, and single colonies were passaged for 3 days on selection plates before being confirmed by DNA sequencing of PCR amplified purified genomic DNA.

Site-directed mutagenesis of VapC-1. VapC-1 D6N and D99N mutants cloned in tandem with a wild-type VapB-1 in the vector pET24b (Millipore Sigma, Burlington, MA USA) were previously described (84). These constructs were used as templates for plasmid expression sets. Synthetic constructs of VapC-1 with E43Q, E120Q, and E43Q D99N mutations were generated in tandem with wild-type VapB-1 by Eurofins (Eurofins Genomics, Louisville, KY USA). The VapC-1 D6N D99N construct, also in tandem with the wild-type VapB-1, was produced by Genscript (Piscataway, NJ USA). Each construct was confirmed by DNA sequencing (Eurofins Genomics, Louisville, KY USA) following sub-cloning into the vectors listed in Table 3.

TABLE 3 Bacteria, plasmids and primers used in this study

Strains	**Description**	**Source**
DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Laboratory collection
LMG194	F$^-$ Δ*lacX*74 *galE thi rpsL* Δ*phoA* (*PvuII*) Δ*ara*714 *leu*::Tn10	Invitrogen
SU101	*lexA*71::Tn*5*(Def) *sulA211* Δ(*lacIPOZYA*)*169*/F' *lacI*^{q} *lacZ M15*::Tn*9*	M. Granger-Schnarr
SU202	*lexA*71::Tn*5*(Def) *sulA*211 Δ(*lacIPOZYA*)169/F' *lacI*^{q} *lacZ* M15::Tn9	M. Granger-Schnarr
86-028NP Δ*vapBC-1*	Strain 86-028NP with the *vapBC-1* locus deleted	(36)
Plasmids	**Description**	**Source**
pBAD33	Highly regulated expression vector	(77)
pET24b	Regulated expression vector with His tag	Novagen
pBluescript SK(+)	Cloning vector	(36)
pSR658	WT LexA DBD fusion vector	(74)
pSR659	Mutant LexA DBD fusion vector	(74)
pSR660	WT LexA DBD fusion vector	(74)
pSR661	Mutant LexA DBD fusion vector	(74)
pDD686	*vapBC-1* in pET24b	(46)
pDD757	*vapBC-1* (*vapC-1* D6N*)* in pET24b	Unpublished
pDD758	*vapBC-1* (*vapC-1* D99N) in pET24b	(84)
pDD866	*vapB-1* in pSR658	(36)
pDD935	*vapB-1* in pTrcHisA	Unpublished
pDD946	*vapC-1* in pBAD33	Unpublished
pDD1058	5' end of NTHI_RS09270 in pBluescript SK(+)	This work
pDD1063	3' end of NTHI_RS09270 in pDD1058	This work
pDD1118	*vapBC-1* (*vapC-1* D6N D99N) in pUC57	Genscript
pDD1119	*vapC-1* D6N D99N in pBAD33	This work
pDD1120	*vapC-1* D99N in pSR658	This work

TABLE 3 Continued

Plasmids	Description	Source
pDD1123	*vapC-1* D99N in pSR659	This work
pDD1126	*vapC-1* D6N in pSR661	This work
pDD1128	*vapC-1* D6N D99N in pSR661	This work
pDD1132	*vapC-1* D6N in pBAD33	This work
pDD1133	*vapC-1* D99N in pBAD33	This work
pDD1140	NTHI_RS09270-targeted delivery vector	This work
pDD1150	*vapC-1* in pSR659	This work
pDD1152	*vapC-1* D6N in pET24b	This work
pDD1153	*vapC-1* D99N in pET24b	This work
pDD1154	*vapC-1* D6N D99N in pET24b	This work
pDD1155	pDD1140 with the native *vapBC-1* promoter	This work
pDD1159	pDD1155 with *vapB-1vapC-1* D6N	This work
pDD1160	pDD1155 with *vapB-1vapC-1* D99N	This work
pDD1161	pDD1155 with *vapB-1vapC-1* D6N D99N	This work
pDD1165	*vapBC-1* (*vapC-1* E43Q) in pCR2.1	Eurofins
pDD1166	*vapBC-1* (*vapC-1* E120Q) in pCR2.1	Eurofins
pDD1167	pDD1155 with *vapB-1vapC-1* E43Q	This work
pDD1168	pDD1155 with *vapB-1vapC-1* E120Q	This work
pDD1169	pDD1155 with *vapB-1vapC-1* wild-type	This work
pDD1170	*vapC-1* E43Q in pSR658	This work
pDD1173	*vapC-1* E43Q in pET24b	This work
pDD1174	*vapC-1* E120Q in pBAD33	This work
pDD1175	*vapC-1* E43Q in pBAD33	This work
pDD1177	*vapC-1* E120Q in pSR658	This work
pDD1178	*vapC-1* E43Q in pSR659	This work
pDD1179	*vapC-1* E120Q in pSR659	This work
pDD1194	*vapC-1* E43Q D99N in pBAD33	This work
pDD1196	*vapC-1* E43Q D99N in pBAD/*Myc*-HisA	This work

TABLE 3 Continued

Plasmids	**Description**	**Source**
pDD1199	*vapC-1* E43Q D99N in pSR658	This work
pDD1200	*vapC-1* E43Q D99N in pSR659	This work
pDD1204	pDD1155 with *vapB-1vapC-1* E43Q D99N	This work
Primers	**Description**	**Source**
D6NFor	[Phos]AACACCAATATCATTATTTATTTAATG	IDT
D6NRev	[Phos]TAACATATAAATCATAAATTTTCTCG	IDT
321LexFor	GAGGGAGCTCATGCTTACTAAAGTG	IDT
321LexRev	AACAGGTACCTCATAAATTTTCTCG	IDT
322LexFor	GAGAGAGCTCATGATTTATATGTTAG	IDT
322LexRev	TCAGGGTACCCTATTTTGTCCAATCTTGCC	IDT
pETD6NC1SacFor	AAAAGAGCTCTATGATTTATATGTTAAACACC	Eurofins
322Rev	TCAGAAGCTTCTATTTTGTCCAATCTTGCC	IDT
pBAD86C1SacFor	TTACGAGCTCAGGAGCGAGAAAATTTATG	Eurofins
XLacIFor	GTCATCTAGAAAACGCGCGAGGCAGC	Eurofins
XHisRev	TATTTGTCTAGAGGCAGTTCCCTACTCTCG	Eurofins
86C1NdeD6NFor	AAACATATGATTTATATGTTAAACAC	Eurofins
VapCXhoRev	GAATCTCGAGTTTTGTCCAATCTTGCC	Eurofins
86C1NdeFor	AAAACATATGATTTATATGTTAGAC	Eurofins
PermCSpeFor	AAATACTAGTAACACACACGCCATTCC	Eurofins
ErmCSpeRev	GTTAACTAGTGCAGTTATGCATCC	Eurofins
BC1PromBamFor	ACTAGGATCCATCATTTACTCATTGACTTGC	Eurofins
BC1PromNdeRev	TAAGCATATGTACCCTCTCGTATATAC	Eurofins
86B1NdeFor	GGTACATATGCTTACTAAAGTG	Eurofins
2009-1SacFor	TGCCGAGCTCGCAGGTAAGATGTCG	Eurofins
2009-1XbaRev	AAAATCTAGAGATGCACGTGCAACC	Eurofins
2009-2XhoFor	AAAACTCGAGGCTTGCATCTAATGC	Eurofins
2009-2KpnRev	AAAAGGTACCTCACTATCTTGTGCC	Eurofins

Purification of VapC-1 for crystallization studies. To overexpress VapBC-1 for purification, pDD686 in BL21(DE3) (46) was grown to logarithmic phase at 37 °C with shaking, induced for 18 h with 1 mM IPTG at 26 °C and harvested at 2,500 x g into pellets from 25 mL aliquots. After one freeze-thaw cycle, VapBC-1 was isolated in accordance with the MagneHis protein purification system (Promega, Madison, WI USA) with the following exceptions: the lysis solution was 1X BugBuster protein extraction reagent (Millipore Sigma, Burlington, MA USA) with 100 mM HEPES pH 7.2, 300 mM NaCl, 1X ProteCEASE-50 (EDTA-free), and 2 µL/mL RNase-free DNase (Thermo Fisher Scientific, Waltman, MA USA); the wash buffer contained 100 mM HEPES pH 7.2 , 300 mM NaCl and 20 mM imidazole, and the elution buffer contained 100 mM HEPES pH 7.2, 300 mM NaCl and 500 mM imidazole. 1 µg of purified VapBC-1 was separated on a Bolt precast 4-12% polyacrylamide gel and analyzed via Western blot using the iBlot and iBind system (Thermo Fisher Scientific, Waltman, MA, USA). The nitrocellulose blot was probed with a monoclonal anti-His tag primary antibody (THE anti-His, Genscript, Piscataway, NJ USA), goat anti-mouse IgG-HRP secondary antibody (Genscript, Piscataway, NJ USA), detected with SuperSignal West Dura chemiluminescence substrate (Thermo Fisher Scientific, Waltman, MA, USA) and imaged on a UVP BioSpectrum 815 (UVP LLC., Upland, CA USA).

Crystallization and data collection. A purified preparation of the NTHi VapBC-1 protein complex was concentrated to 20 mg/mL in 20 mM NaCl, 10 mM HEPES pH 7.2 for crystallization screening. All crystallization experiments were set up using an NT8 drop setting robot (Formulatrix Inc.) and UVXPO MRC (Molecular Dimensions) sitting drop vapor diffusion plates at 18 °C. 100 nL of protein and 100 nL of crystallization solution

were dispensed and equilibrated against 50 µL of the latter. Crystals approximately 100 µm long that displayed a needle morphology were observed within one week from the Proplex screen (Molecular Dimensions) condition C3 (20% [wt/vol] PEG 4000, 100 mM sodium acetate pH 5.0, 200 mM ammonium acetate). A cryoprotectant solution composed of 80% crystallization solution and 20% (vol/vol) PEG 200 was dispensed (2 uL) onto the drop, crystals were harvested with a cryoloop immediately and stored in liquid nitrogen. X-ray diffraction data were collected at the Advanced Photon Source beamline 17-ID using a Dectris Pilatus 6M pixel array detector.

Structure solution and refinement. Intensities were integrated using XDS (90, 91) via Autoproc (92) and the Laue class analysis and data scaling, using two data sets to increase the multiplicity, were performed with Aimless (93) which indicated that the crystals belong to the *mmm* Laue class. Structure solution was conducted by molecular replacement with Morda (94) to place a VapC-1 dimer using *Shigella flexneri* VapC (PDB 5ECD) as the template and the top solution was obtained in the space group $P2_12_12_1$. The VapC-1 coordinates from Morda were input as a fixed solution into Phaser (95) and a VapB monomer from the *Rickettsia felis* VapBC structure (PDB 3ZVK) was used as a search model. The final solution from Phaser consisted of a VapB-1 and VapC-1 dimer in the asymmetric unit and the model was improved by automated building with Arp/wARP (96). Further refinement and manual model building were conducted with Phenix (97) and Coot (98), respectively. Disordered side chains were truncated to the point for which electron density could be observed. Structure validation was conducted with Molprobity (99) and figures were prepared using the CCP4MG package (100). Structure superposition was carried out with GESAMT (101). Crystallographic data are provided in

Table 4. Coordinates and structure factors were deposited to the worldwide Protein Data Bank (wwPDB) with the accession code 6NKL.

TABLE 4 Crystallographic data for VapBC-1

Parameter	Value(s) for VapBC-1
Data Collection	
Unit-cell parameters (Å, °)	*a*=43.88, *b*=57.32, *c*=175.75
Space group	$P2_12_12_1$
Resolution (Å)[a]	48.01-2.20 (2.28-2.20)
Wavelength (Å)	1.0000
Temperature (K)	100
Observed reflections	303,859
Unique reflections	23,388
$\langle I/\sigma(I)\rangle$[a]	10.9 (2.1)
Completeness (%)[a]	100 (100)
Multiplicity[a]	13.0 (13.7)
R_{merge} (%)[a,b]	16.8 (161.3)
R_{meas} (%)[a,d]	17.5 (167.7)
R_{pim} (%)[a,d]	4.9 (45.2)
$CC_{1/2}$ [a,e]	0.998 (0.784)
Refinement	
Resolution (Å)[a]	48.01-2.20
Reflections (working/test)[a]	22,146/1,166
R_{factor} / R_{free} (%)[a,c]	18.6/23.4
No. of atoms (VapB/VapC-1/Water)	1,085 / 2,103 / 58
Model Quality	
R.m.s deviations	
Bond lengths (Å)	0.009
Bond angles (°)	0.981
Average *B*-factor (Å^2)	
All Atoms	43.5

TABLE 4 Continued

Parameter	Value(s) for VapBC-1
Data Collection	
VapB-1	48.0
VapC-1	41.3
Water	41.8
Coordinate error	0.24
(maximum likelihood) (Å)	
Ramachandran Plot	
Most favored (%)	96.8
Additionally allowed (%)	3.0

[a]Values in parenthesis are for the highest resolution shell.
[b]$R_{merge} = \Sigma_{hkl}\Sigma_i |I_i(hkl) - \langle I(hkl)\rangle| / \Sigma_{hkl}\Sigma_i I_i(hkl)$, where $I_i(hkl)$ is the intensity measured for the *i*th reflection and $\langle I(hkl)\rangle$ is the average intensity of all reflections with indices *hkl*.
[c]$R_{factor} = \Sigma_{hkl} ||F_{obs}(hkl)| - |F_{calc}(hkl)|| / \Sigma_{hkl} |F_{obs}(hkl)|$; R_{free} is calculated in an identical manner using 5% of randomly selected reflections that were not included in the refinement.
[d]R_{meas}, redundancy-independent (multiplicity-weighted) R_{merge} (93, 102); R_{pim}, precision-indicating (multiplicity-weighted) R_{merge} (103).
[e]$CC_{1/2}$ is the correlation coefficient of the mean intensities between two random half-sets of data (104, 105).

Cloning and protein-protein interactions of VapC-1 mutants. The VapC-1 D99N, E43Q, E43Q D99N, and E120Q mutants were cloned by amplifying the *vapC-1* D99N gene from pDD758, the *vapC-1* E43Q gene from pDD1165, the *vapC-1* E43Q D99N gene from pDD1196, or the *vapC-1* E120Q gene from pDD1166 using high-fidelity Phusion Flash DNA Polymerase (Thermo Fisher Scientific, Waltham, MA USA) and the primers 322LexFor and 322LexRev. The PCR products, pSR658, and pSR659 were digested with *SacI* and *KpnI* prior to ligation, resulting in pDD1120, pDD1123, pDD1170, pDD1178 pDD1199, pDD1200, pDD1177, and pDD1179. VapC-1 D6N and D6N D99N were cloned by amplifying the *vapC-1* D6N gene from pDD757 and the *vapC-1* D6N D99N gene from pDD1118, using high-fidelity Phusion Flash DNA Polymerase with pETD6NC1SacFor and 322LexRev primers. The PCR products, pSR660, and pSR661 were digested with *SacI* and *KpnI* prior to ligation, resulting in pDD1125, pDD1126, pDD1129, and pDD1128. The wild-type VapC-1 toxin was cloned into pSR659 with the above method from NTHi strain 86-028NP genomic DNA (pDD1150). All clones were confirmed by DNA sequencing. The pSR659 or pSR661 constructs were co-transformed with pDD866 into SU202 to perform heterodimerization assays. The pSR658 or pSR660 constructs were individually transformed into SU101 for homodimerization assays.

Expression of the LexA fusions were induced by IPTG and if the fusion proteins interacted, a competent repressor was formed in the reporter strain that could bind the operator site and decrease transcription of the *lacZ* reporter gene. This resulted in a measurable decrease in β-galactosidase activity compared to a strain carrying the control unfused vectors. However, because β-galactosidase is a long-lived enzyme, the reporter strains were grown overnight in the presence of 1 mM IPTG so that any enzyme

expressed prior to the LexA chimera induction is degraded. This strategy results in a more reliable and accurate quantification of homodimerization and heterodimerization. Following overnight incubation on LB agar plates with 1 mM IPTG and appropriate antibiotics, reporter strains carrying the control vectors (pSR658 and pSR659 or pSR660 and pSR661) (75) or the LexA DBD fusions were inoculated and grown to an optical density at 600 nm (OD_{600}) of ~0.5 in LB broth with 1 mM IPTG and appropriate antibiotics. Protein-protein interactions were quantified by β-galactosidase activity assays and compared to the activity of the reporter strain carrying the unfused vectors, which has the highest activity due to the lack of repressor formation. The algorithm for quantifying β-galactosidase activity was: $[OD_{420}-(1.75 \times OD_{550})/(t \times v \times OD_{600})] \times 1000$, where *t* is time of reaction development in minutes and *v* is volume of sample in milliliters (76). This equation allows for the normalization of different culture densities for comparison purposes.

Cloning and growth-recovery assays with VapC-1. To clone wild-type and mutant VapC-1 genes for the growth-recovery assays, the wild-type *vapC-1* gene was amplified from NTHi strain 86-028NP genomic DNA, the *vapC-1* D6N gene was amplified from pDD757, the *vapC-1* D99N gene was amplified from pDD758, the *vapC-1* D6N D99N gene was amplified from pDD1118, the *vapC-1* E43Q gene was amplified from pDD1165, the *vapC-1* E43Q D99N gene was amplified from a synthetic gene produced by Eurofins, and the *vapC-1* E120Q gene was amplified from pDD1166 using high-fidelity Phusion Flash DNA polymerase with pBAD86C1SacFor and 322Rev primers. The PCR products and pBAD33 were digested with *SacI* and *HindIII* prior to ligation, resulting in pDD946, pDD1132, pDD1133, pDD1119, pDD1175 pDD1194, and pDD1174. The forward primer

was designed to include the putative Shine-Dalgarno sequence for VapC-1. The *vapB-1* gene was amplified from 86-028NP genomic DNA using high-fidelity Phusion Flash DNA Polymerase with 321LexFor and 321LexRev primers. After digestion with *SacI* and *KpnI* the PCR fragment was ligated into pTrcHisA, resulting in pDD935. All clones were verified by DNA sequencing. Wild-type *vapC-1* and each of the six mutants were co-transformed with pDD935 into *E. coli* strain LMG194.

To perform the growth-recovery assays, strains were grown at 37 °C to early log phase (OD_{600} of ~0.2 - 0.3) in LB broth with 20 µg/mL chloramphenicol and 100 µg/mL ampicillin before being diluted in a 1:1 (vol:vol) ratio to a final volume of 150 microliters of media in a 96-well Corning™ Falcon™ polystyrene microplate (Corning, Corning, NY, USA) containing LB broth with 20 µg/mL chloramphenicol and 100 µg/mL ampicillin, with or without 1 mM IPTG, 6.6 mM (0.1% wt/vol) arabinose, or both. The final induction concentrations were 0.5 mM IPTG and 3.3 mM (0.05% wt/vol) arabinose. The plate was covered with a gas permeable membrane, loaded into a Multiskan FC incubating plate reader (Thermo Fisher Scientific, Waltham, MA USA), and incubated at 37 °C with a shaking protocol of 5 *s* on and 5 *s* off with measurements taken every 10 min at OD_{595} for 190 min. Each assay included three biological replicates with three technical replicates.

Construction of the *in cis* delivery system and reconstituted NTHi strains. Based on a previously published ectopic delivery vector in NTHi (106), the *metE* pseudogene in the chromosome of strain 86-028NP (NTHI_RS09270) was targeted as the site for gene delivery and single-copy expression *in cis*. The delivery vector was constructed by amplifying 843 bp of the 5’ end of NTHI_RS09270 from 86-028NP genomic DNA using high-fidelity Phusion Flash DNA polymerase with 2009-1SacFor and

2009-1XbaRev primers and ligating the product into pBluescript SK$^+$ digested with *SacI and XbaI*, resulting in pDD1058. The second homologous region of 767 bp of the 3' end of NTHI_RS09270 was amplified from 86-028NP genomic DNA using high-fidelity Phusion Flash DNA polymerase with 2009-2XhoFor and 2009-2KpnRev primers and ligated into pDD1058 cut with *XhoI* and *KpnI*, resulting in pDD1063. The last step was amplifying a erythromycin resistance cassette from pEJ18 (107) with PermCSpeFor and ErmCSpeRev primers and ligating into pDD1063 cut with *SpeI*, resulting in the delivery vector, pDD1140. The native *vapBC-1* promoter from 86-028NP was cloned into the multiple cloning site of the delivery vector using the primers BC1PromBamFor and BC1PromNdeRev. The reverse primer contains an engineered *NdeI* site (CATATG) in which the ATG of the restriction site (underlined) serves as the start codon for the fused *vapBC-1* genes. This allowed us to clone all constructs with their native transcriptional regulation and organization. The *vapBC-1* promoter was ligated to pDD1140 cut with *BamHI and EcoRV*, resulting in pDD1155. Because the *vapC-1* mutants are in tandem with the wild-type *vapB-1* gene, each was cloned into pDD1155 for delivery into the 86-028NP chromosome. The *vapBC-1* D6N locus was amplified from pDD757, *vapBC-1* D99N was amplified from pDD758, *vapBC-1* D6N D99N was amplified from pDD1118, *vapBC-1* E43Q was amplified from pDD1165, *vapBC-1* E43Q D99N was amplified from a synthetic gene produced by Eurofins, *vapBC-1* E120Q was amplified from pDD1166, and wild-type *vapBC-1* was amplified from pDD686 using high-fidelity Phusion Flash DNA polymerase with 86B1NdeFor and 322Rev primers. Each amplicon was ligated to pDD1155 digested with *NdeI* and *HindIII*, resulting in pDD1159, pDD1160, pDD1161,

pDD1167, pDD1204, pDD1168, and pDD1169. This resulted in *vapBC-1* with the desired mutations in *vapC-1* under the control of its native promoter (Fig. 12).

The strains of NTHi 86-028NP Δ*vapBC-1* NTHI_RS09270::$P_{vapBC-1}$::*vapB-1vapC-1* with wild-type *vapC-1* and the six mutants were constructed by homologous recombination into the chromosome of 86-028NP Δ*vapBC-1* by PCR products containing both the flanking regions of NTHI_RS09270 and each promoter::tandem clone using MIV transformation (81). After selection on chocolate agar with 5 µg/mL erythromycin, a single colony of each transformant was passaged three times on antibiotic-containing plates. DNA sequencing of PCR-amplified genomic DNA preparations confirmed each correct strain.

To ensure the *in cis* fusions did not affect the growth dynamics of NTHi, strains were grown to early log phase (OD_{600} of ~0.2 - 0.3) in sBHI broth before dilution into a 96-well Corning™ Falcon™ polystyrene microplate, and absorbance was measured following the same protocol as the growth-recovery assays over the course of 7 h. Each growth assay included three biological replicates with three technical replicates.

NTHi infections of primary human tissues. EpiAirway™ tissues (AIR-100-ABF, MatTek, Ashland, MA USA) were maintained, inoculated, and harvested as previously described (108) with the following exception: inserts were inoculated with approximately 25 µl 1.0 x 10^4 CFU/mL bacterial suspension of the desired strain. Inserts were cocultured with the parent, wild-type or mutant strains for 48 h before harvesting and quantification by counting viable colonies from a plated dilution series on chocolate agar plates.

Statistical analyses. Differences between means were determined using the Student's *t*-test. Differences among multiple-group treatments were determined by the

repeated measures ANOVA or two-tailed ANOVA as appropriate using GraphPad Prism 8.0.1 (GraphPad Software, La Jolla, CA USA). A p value of ≤ 0.05 was considered statistically significant.

Acknowledgments. This work was supported in part by the National Institute on Deafness and other Communication Disorders cooperative agreement U01 DC014756 to DAD. The authors thank L.C. Zavada for technical support. The National Institute on Deafness and other Communication Disorders (NIDCD) had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results. NPC and MS were supported by the National Center for Advancing Translational Sciences (NCATS) Division of Pre-Clinical Innovation Intramural Program. Use of the University of Kansas Protein Structure Laboratory was supported by a grant the National Institute of General Medical Sciences (P30 GM110761) at the National Institutes of Health. Use of the IMCA-CAT beamline 17-ID at the Advanced Photon Source was supported by the companies of the Industrial Macromolecular Crystallography Association through a contract with Hauptman-Woodward Medical Research Institute. Use of the Advanced Photon Source was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. This research is in accordance with protocols 17-007 and 15-006 approved by the Institutional Biosafety Committee (IBC), Old Dominion University, USA.

USE OF THE L-ARABINOSE-INDUCIBLE *ESCHERICHIA COLI ARABAD* PROMOTER AND THE *ARAC* REGULATOR IN *HAEMOPHILUS INFLUENZAE*

INTRODUCTION

The discovery of inducible promoters to express genes has provided scientists with valuable tools to study gene expression. There are a wide variety of promoters with various regulators and induction methods available; however, the P_{trc} (78), P_{lac} (109-111), P_{tac} (111-113), P_{T7} (114), and P_{BAD} (77) promoters are among the more popular expression systems. The P_{BAD} expression vectors are extremely useful due to the tight regulation of the promoter via the regulatory gene, *araC*, the fast induction rate, and the ability to modulate the level of gene expression (77).

The arabinose system has been studied extensively in *E. coli* and is composed of two independent transport systems (*araE* and *araFGH*), a set of metabolism genes known as the *araBAD* operon, and the regulatory gene, *araC*. The product of *araE* is a low-affinity membrane protein that uses the electrochemical potential to transport arabinose into the cell (115, 116). The *araFGH* operon is a second high-affinity arabinose transporter belonging to the ATP-binding cassette (ABC) transporter family (117, 118). AraF is the arabinose-binding protein located in the periplasm (119), AraG binds ATP, and AraH is a membrane-associated protein involved in arabinose transport into the cell (117, 120). The *araBAD* operon converts L-arabinose to D-xylulose-phosphate, part of the pentose phosphate pathway, for use by bacteria in central metabolism (121). AraC is the regulatory protein of the arabinose system (122, 123). It positively regulates the promoters of *araBAD, araE, and araFGH* in the presence of arabinose (124); it negatively regulates the *araBAD* promoter in the absence of arabinose, and it negatively regulates

its own promoter in the presence and absence of arabinose (125). Two mechanisms describe the regulatory functions of the arabinose-specific AraC protein: DNA looping (126, 127) and the light-switch mechanism (128, 129). In DNA looping, a protein or protein complex binds to two different sites of DNA causing it to create a loop (130). This looping decreases the concentration of the regulator, for example AraC, while still repressing gene expression. Conversely, the light-switch mechanism facilitates positions and orientations of the regulator protein(s) that favors DNA looping (131). When the regulator protein is in one position, DNA looping occurs and when the interaction changes, the loop can be broken to initiation transcription, simulating the on and off function of a light-switch. AraC functions as a homodimer. Each monomer has two domains connected by a flexible linker: a dimerization domain that binds arabinose and a DNA-binding domain (Fig. 14). In the absence of arabinose, AraC binds to $araO_2$ (O_2) and araI (I_1) causing the DNA to bend, creating a loop (126, 127). Meanwhile, the N-terminal arm extends from the arabinose-binding domain to the C-terminal DNA-binding domain, keeping the arabinose binding site exposed for arabinose when it becomes available (132). When arabinose arabinose is present, it sits in the arabinose-binding domain making it energetically favorable for the N-terminal arms to fold over the pocket and bind to the dimerization domains (129). When the arms release from the DNA-binding domains, it is energetically favorable for AraC to bind to adjacent I_1 and I_2 half-sites, providing space for RNA polymerase to bind to the P_{BAD} promoter to initiate transcription (133, 134). The full induction of the arabinose genes requires AraC as well as the catabolite activator protein (CAP; also known as the cyclic AMP receptor protein) to bind upstream of the I_1 and I_2

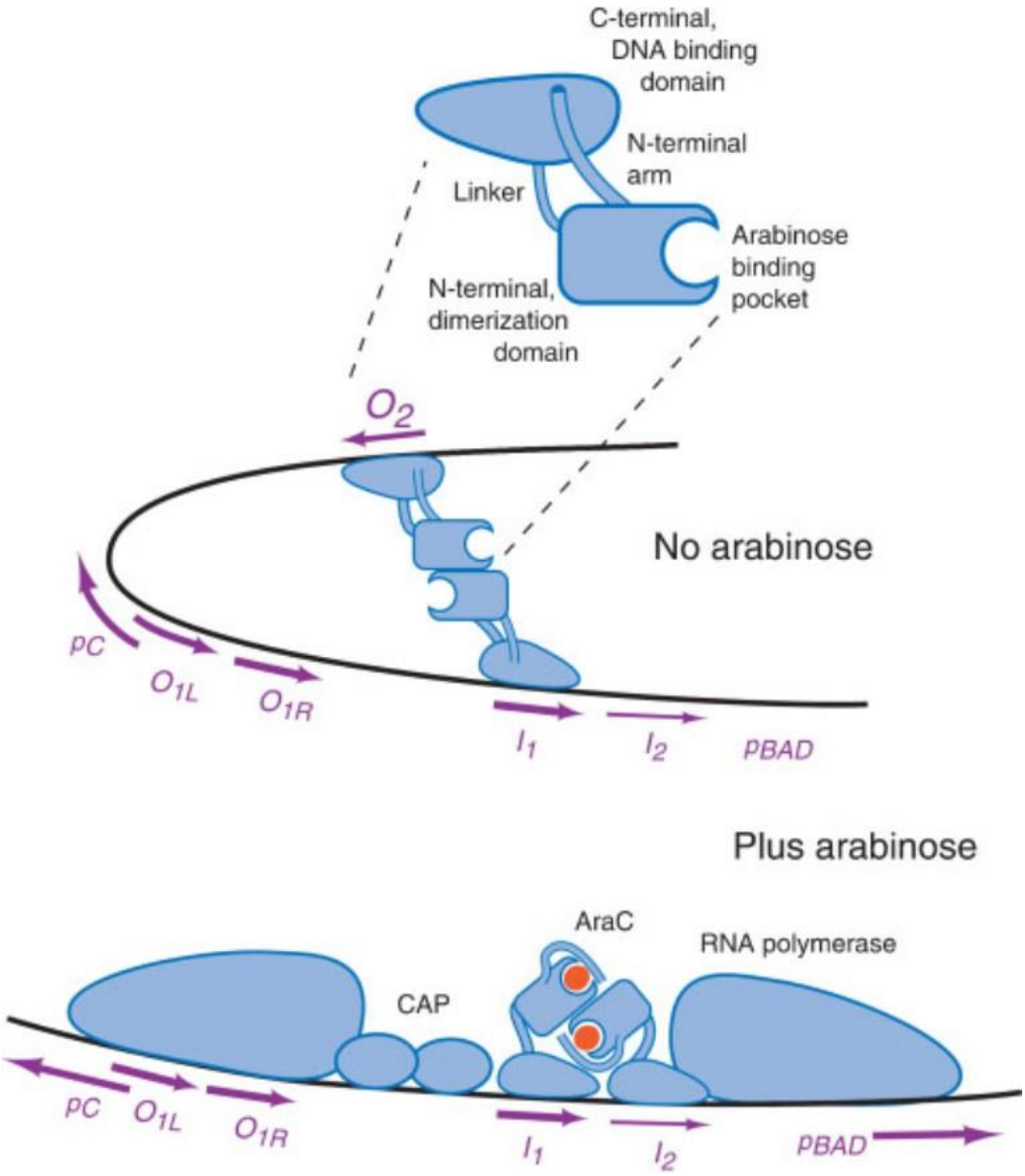

FIG 14 Regulation of the P_{BAD} promoter by AraC. AraC is a homodimer. The C-terminal DNA binding domains bind to the O_2 and I_1 sites when arabinose is not present. This energetically favorable binding causes DNA to form a loop and inhibits RNA polymerase from binding to the P_{BAD} and P_C promoters. The N-terminal dimerization domains have arabinose-binding pockets. When arabinose binds, the N-terminal arm which was bound to the DNA binding domain folds over the arabinose-binding domain in a process called the light-switch mechanism. This facilitates a conformation change in the AraC homodimer and the DNA-binding domain that was bound to the O_2 site folds over to bind to the I_2 site. The DNA loop is released and RNA polymerase can now bind to the P_{BAD} and P_C promoters to facilitate transcription. This image is reproduced from reference 132 describing the history and mechanisms of AraC.

operator sites (Fig. 14) and stimulate the transcription of the *araBAD* operon (77, 135). Induction of the arabinose genes is inhibited by glucose levels because glucose reduces levels of cAMP by inhibiting adenylate cyclase.

The expression system was originally designed for use in *E. coli* (77). However, the need to achieve high levels of gene expression, particularly of toxic gene products, has enticed scientists to expand the practical nature of the arabinose system into other bacterial species. Scientists have utilized the P_{BAD} promoter and the *araC* regulator for routine cloning and gene expression to facilitate protein purification, analyze protein function, and study complex cellular physiological properties in the following species: *Salmonella typhimurium (136), Agrobacterium tumefaciens* (137), *Corynebacterium glutamicum* (138), *Xanthomonas campestris* pv. *phaseoli (139)*, *Pseudomonas aeruginosa* (140), *Vibrio cholerae* (141), and *Burkholderia cepacia* (142).

Previous studies have found the P_{BAD} expression vectors to be extremely useful to study the ribonuclease toxin from type II TA loci. A growth recovery assay which utilized the plasmid pBAD33 (77) allowed induction of gene expression by arabinose to evaluate ribonuclease activity of VapC-1 mutants from NTHi in *E. coli* LMG194 (143).

While it is common to utilize *E. coli* to study genes expressed in other bacterial species, in this study a novel assay was designed in the background of NTHi to test both the transport of arabinose and the ability of this molecule to regulate the expression of TA proteins *in cis.* Ideally, to study arabinose induction, a bacterial strain should be able to transport but not degrade arabinose, so that the molecule remains available to induce the P_{BAD} promoter. An extensive BLAST search of NTHi genomes did not reveal genes that were homologous to known arabinose transport proteins and catabolic enzymes (144). In

this study, two novel NTHi delivery vectors were constructed: one that carried an *araC* gene with a P_{BAD}*::vapD* expression cassette, and another that delivers an arabinose permease into the NTHi chromosome. Both genes were inducible by arabinose. These were used to transform an NTHi strain with a deletion of the TA locus *vapXD* (R2866 Δ*vapXD)*. Because the cassettes were flanked by DNA regions homologous to a target site on the chromosome and included an antibiotic marker (Fig. 16 and 17), they were recombined at the desired site following transformation and identification was facilitated by antibiotic selection followed by DNA sequencing of the *in cis* cassette. Using this novel design, the ability of the *araC* P_{BAD}*::vapD* fusion to recombine into the NTHi chromosome at the site of a naturally-occurring pseudogene was evaluated. Then, one of two *E. coli* arabinose permease cassettes (comprised of either *araE* or *araFGH*) was targeted to another non-utilized site on the chromosome, a partial bacteriophage terminase gene. Upon successful homologous recombination of the toxin only, or the toxin with one of the two arabinose permease cassettes, expression of the toxin VapD with or without an arabinose permease was induced. Likewise, the ability of each permease to transport arabinose when expressed in NTHi was compared. A bioassay was then designed to measure the amount of arabinose uptake by NTHi by detecting the amount of arabinose left in the media after NTHi growth (hereinafter called conditioned media). This bioassay uses the expression of a green fluorescent protein (GFP) gene under the control of the P_{BAD} promoter from the plasmid pGLO (Bio-Rad, Hercules, CA) in *E. coli* strain LMG194 which is inducible by arabinose. First, it was demonstrated that arabinose could modulate GFP expression in chemically defined media using a dose-response control curve. Subsequently, this bioassay was used to measure the arabinose remaining in conditioned

media following NTHi growth at mid-log phase and also at stationary phase to determine whether 1) NTHi could transport arabinose across its membrane, and 2) whether the addition of either *E. coli* arabinose permease affected arabinose transport in NTHi.

RESULTS

Growth dynamics of R2866 and the *ΔvapXD* mutant strain. To induce the expression of the toxin VapD and antitoxin VapX in the background of NTHi, it was necessary to create a *vapXD* deletion mutant to prevent crosstalk. The entire *vapXD* operon was deleted in R2866 by replacing the locus with a chloramphenicol resistance cassette. Following confirmation of deletion by PCR characterization of genomic DNA using various flanking primers, growth curves were performed to determine if the deletion had an effect on replication compared to the parent strain. No significant difference in the growth dynamics between the two strains was observed in sBHI (Fig. 15) and in glucose free DM with 10mM sodium pyruvate (data not shown).

Construction of two *in cis* ectopic systems to study type II TA protein induction and arabinose permease function in NTHi. Use of a growth-recovery assay in the background of *E. coli* LMG194 to study the activity of NTHi VapC-1 mutants under the control of the P_{BAD} promoter *in trans* was previously reported (143). While this is an excellent tool, this study sought to create an assay in the background of NTHi to study the function of the TA toxin VapD *in cis* to mimic the single copy toxin gene in the chromosome. In addition, this study also sought to investigate whether the addition of an *E. coli* arabinose permease *in cis* would increase arabinose transport across the NTHi membrane.

To evaluate these aspects, two novel systems were developed. The first contained both the arabinose operon regulatory gene, *araC* and the P_{BAD} promoter controlling the type II toxin, *vapD* (Fig. 16). This system mimicked the normal copy number of the gene and placed it under the control of the artificially-inducible promoter P_{BAD}. This module was

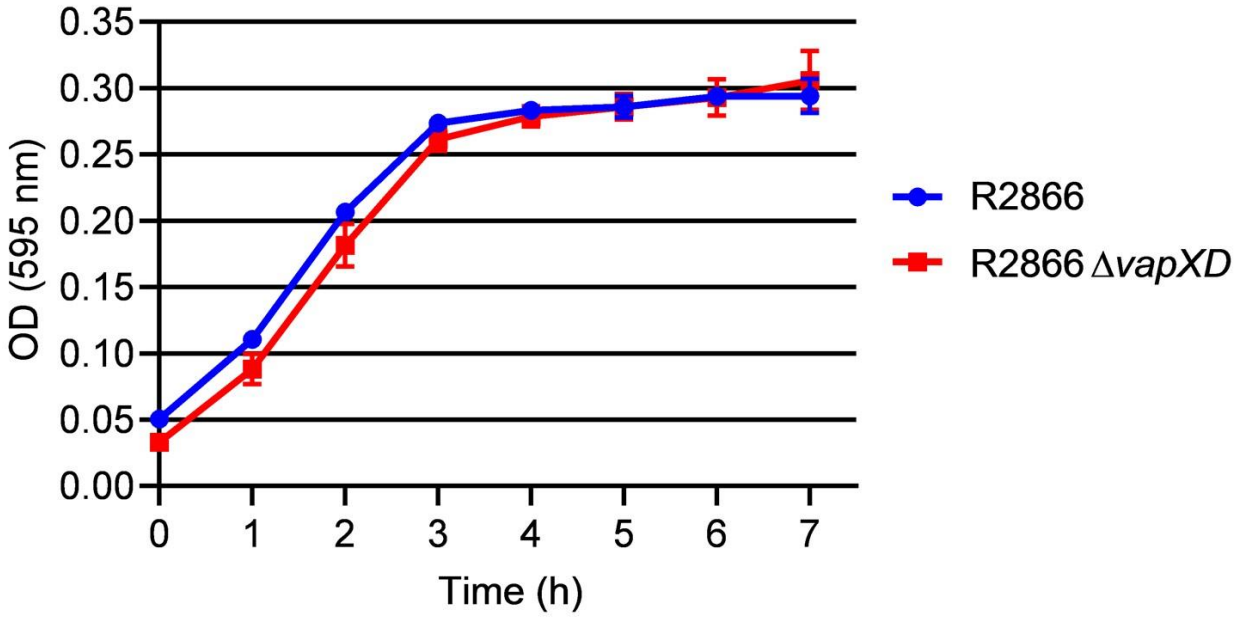

FIG 15 Growth dynamic comparisons between R2866 and *ΔvapXD* deletion mutant strains in sBHI. Strains were grown in a 96-well plate at 35°C with shaking. Data are expressed as the mean ± SD of three biological replicates in triplicate. No significant difference was found between the growth of the parent strain, R2866 and the *ΔvapXD* deletion mutant using the repeated-measures ANOVA.

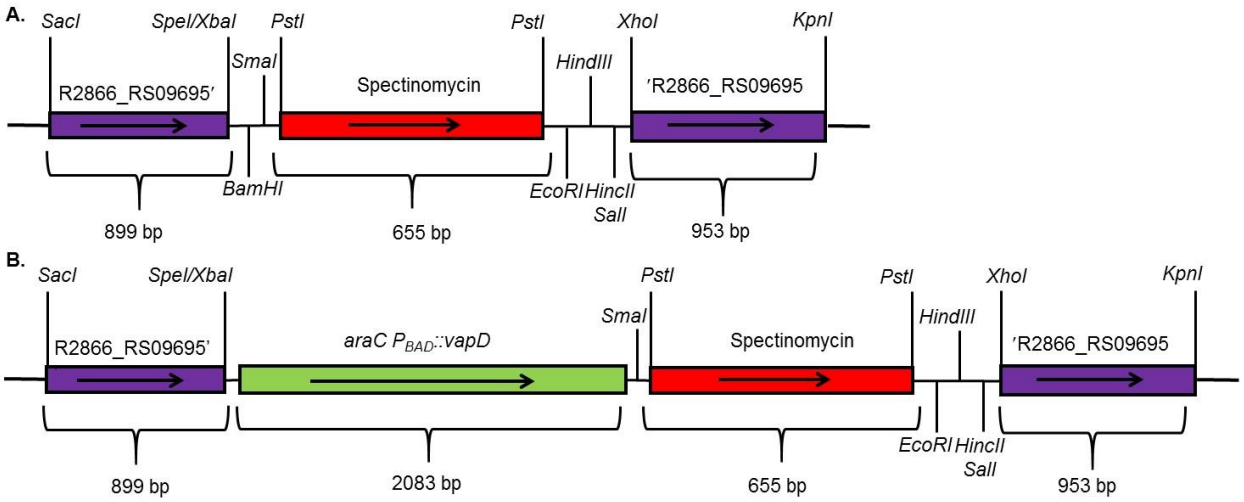

FIG 16 R2866_RS09695 delivery vector for single-copy *in cis* expression of cloned genes. (A) The construct targets the *metH* pseudogene in strain R2866; (B) contains the*araC* gene with the P_{BAD} promoter fused to *vapD.*

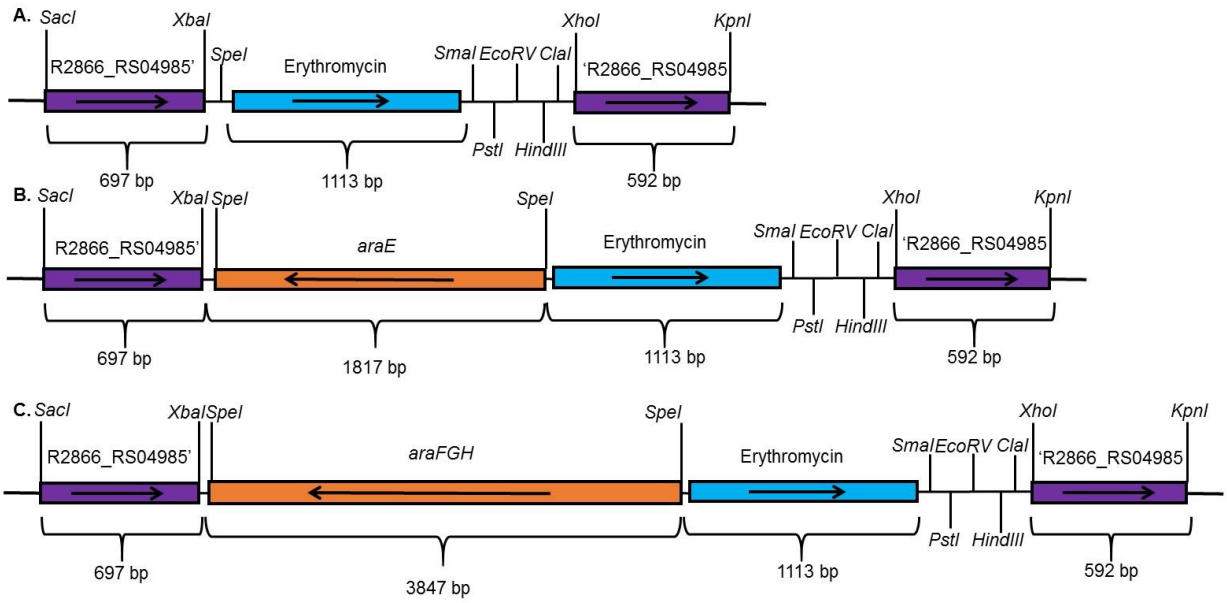

FIG 17 R2866_RS04985 delivery vector for single-copy *in cis* expression of cloned genes. (A) This construct targets the large subunit of the PBSX phage terminase gene in strain R2866; (B) contains the *araE gene*; (C) contains the *araFGH* operon cloned into pDD1230 under the control of their own promoters and are induced by arabinose.

flanked by homologous regions of the *metH* pseudogene from the NTHi strain R2866 (R2866_RS09695), which targets the *araC* P_{BAD}*::vapD* cassette to a chromosomal region that is not in use. Briefly, cassettes were amplified by PCR, used to transform NTHi, and were homologously recombined into the chromosome at the R2866_RS09695 pseudogene site in single copy. An R2866 *ΔvapXD* deletion strain was constructed as the background to analyze *vapD* induction to eliminate contributions by the wild-type *vapXD* locus. The second cassette targeted either the low-affinity arabinose permease gene, *araE*, or the high-affinity arabinose operon, *araFGH* into the large subunit of the PBSX phage terminase gene (R2866_RS04985) (Fig. 17), which was another region not utilized in NTHi. This novel approach resulted in a more biologically relevant system to study the induction of type II ribonuclease toxins under the control of the artificially-inducible P_{BAD} promoter and to determine whether the addition of an *E. coli* arabinose permease increased arabinose transport in the background of NTHi.

NTHi growth recovery assay. Traditionally, the toxin proteins of TA modules induce varying levels of growth arrest, whereas the antitoxins usually do not. Therefore, the NTHi-based growth recovery assay was designed with the toxin in single copy *in cis* and the antitoxin in multicopy *in trans*. The antitoxin, *vapX,* was evaluated *in trans* under the control of the P_{trc} promoter and its expression was induced with 1 mM IPTG (isopropyl-β-D-thiogalactopyranoside) (145, 146). The toxin, *vapD,* was evaluated *in cis* under the control of the P_{BAD} promoter and was induced with 2.64 mM (0.04% wt/vol) arabinose (77) with or without either the low-affinity arabinose permease, *araE,* or the high-affinity arabinose permease, *araFGH in cis* (Fig. 18A to C). Thus, VapX and VapD could be

induced either independently or simultaneously while also comparing the effectiveness of an arabinose permease.

To initiate the assay, a culture in early growth phase was split and treated with either: (i) media only, (ii) arabinose only, (iii) IPTG only, or (iv) both arabinose and IPTG. The results of the experiments with *vapD in cis* with or without either arabinose permease *in cis* induced with arabinose and *vapX in trans* induced with IPTG are presented in Fig. 18A to C. VapD, from NTHi strain 86-028NP, is a ribonuclease toxin (36), and was induced in single copy in the background of NTHi to measure growth arrest. No significant growth arrest was observed when *vapD in cis* was induced *in vivo* with or without the addition of *araE* or *araFGH* (Fig. 18A to 18C, red symbols).

Modulation of expression of GFP in glucose-free defined media. One of the advantages of using the P_{BAD} promoter is the ability to modulate gene expression by varying the arabinose concentration in the media (77). In that context, it is believed that this is the first time an arabinose uptake and utilization system in NTHi defined media has been tested (147). The pGLO (pBAD-GFPuv (148)) plasmid was utilized because the reporter gene, *GFP*, is under the control of the P_{BAD} promoter and therefore is inducible by arabinose. Arabinose in dose-response concentrations was used to induced pGLO in *E. coli* LMG194, using glucose-free defined media with 10 mM sodium pyruvate. Glucose-free media was used because glucose reduces levels of cAMP by inhibiting adenylate cyclase which reduces CAP and therefore reduces induction of the arabinose genes as previously described in the introduction. Therefore, using glucose-free media supplemented with sodium pyruvate can increase the transcription of the *araBAD* operon while still providing NTHi with a suitable carbon source. Relative fluorescence units per

cell (RFU/cell) was calculated from GFP expression measured for each concentration of arabinose (Fig. 19). This experiment shows that arabinose in glucose-free defined media with 10 mM sodium pyruvate can modulate the expression of pGLO in LMG194 (Fig. 19). All treatments were compared and found to be statistically significant using one-way ANOVA.

Transport of arabinose in NTHi. Since the level of expression of GFP from the plasmid pGLO induced in *E. coli* LMG194 reflected the amount of arabinose in the media (Fig. 19), this characteristic was used as a bioassay to determine the concentration of arabinose left in conditioned media after NTHi growth. The control was defined media with arabinose that was treated identically to the conditioned media, but without NTHi. If the amount of arabinose left in the conditioned media resembled the control media without bacteria, then it would indicate NTHi did not transport arabinose into the cell. Likewise, if the amount of arabinose left in the conditioned media was lower than in the control media without cells, then it would suggest NTHi could transport arabinose from the media into the cell. Finally, if the amount of arabinose left in the conditioned media from the strains with one of the arabinose permeases *in cis* was lower than in the conditioned media from strains with *vapD* only, then it could be due to increased arabinose transport by the permease.

Conditioned media was obtained by inducing the following strains with 2.64 mM (0.04% wt/vol) arabinose: R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD,* R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araE}::*araE,* and R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araFGH}::*araFGH* and growing each to mid-log phase (approximate OD_{600} of 0.4) and to stationary phase (approximate OD_{600} of 1.3). Two growth phases

were chosen to determine if arabinose transport was affected by metabolism and/or cell growth in NTHi. The control was media treated with 2.64 mM (0.04% wt/vol) arabinose with no bacteria. The RFU/cell for the control and for each sample of conditioned media was calculated. The RFU/cell of the control was not significantly different from the conditioned media collected during mid-log or stationary phase for all three strains determined by the one-way ANOVA (Fig. 20). This result indicates that 1) arabinose cannot cross the membrane of NTHi, and 2) that the addition of the arabinose permease gene, *araE*, or the arabinose permease operon, *araFGH* from *E. coli* were not advantageous to arabinose transport in NTHi.

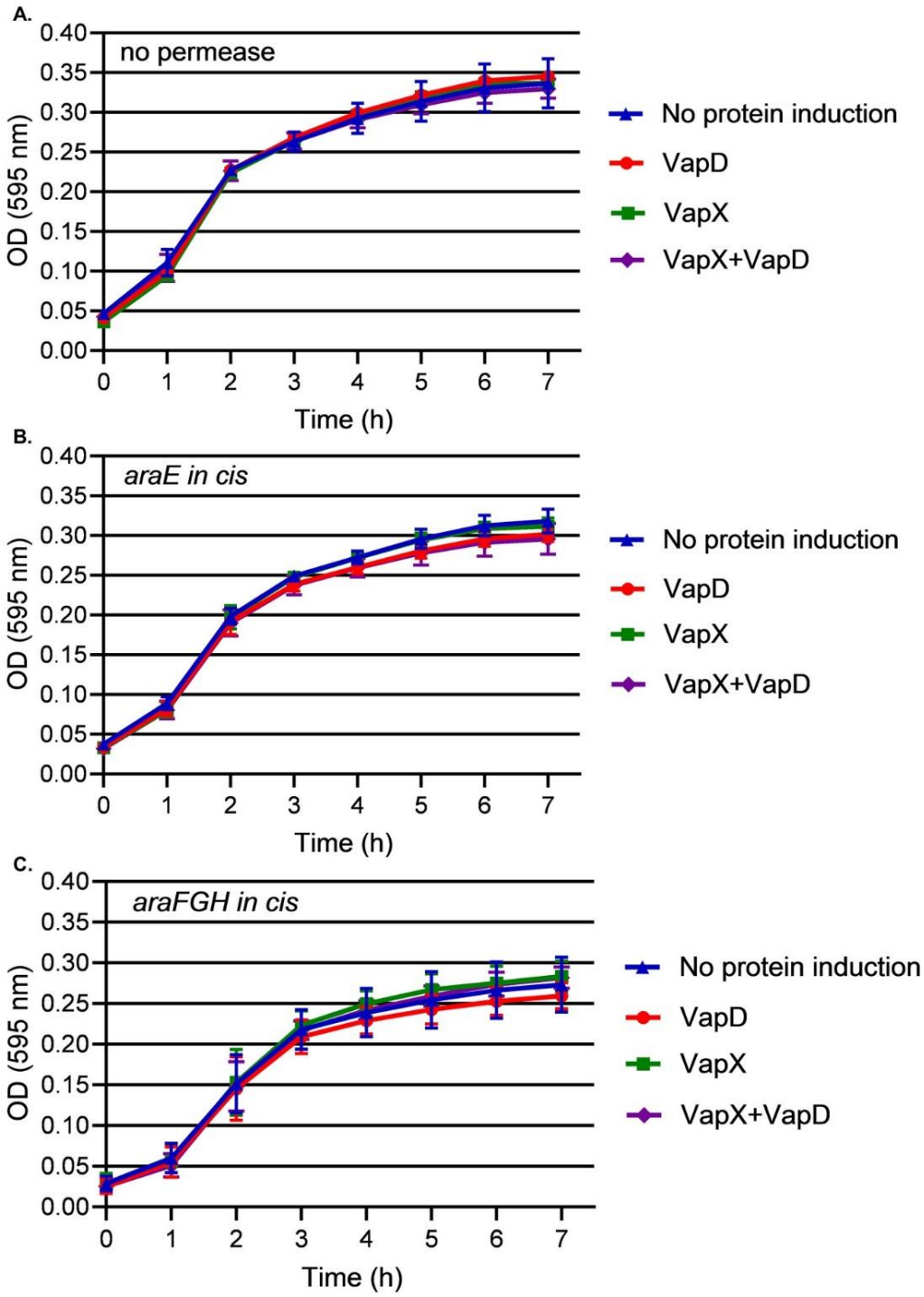

FIG 18 NTHi growth recovery assays in glucose-free defined media with 10 mM sodium pyruvate. NTHi growth-recovery assays with inducible VapX *in trans* and VapD *in cis*. (A) Without an arabinose permease; (B) with *araE in cis*; (C) with *araFGH in cis*. Non-induced NTHi growth (measured by absorbance at 595 nm) or growth following induction of VapD, VapX, or both VapD and VapX. All data are plotted as the mean ± SD of three biological replicates in triplicate. There was no statistical significance among the four treatment groups using repeated-measures ANOVA.

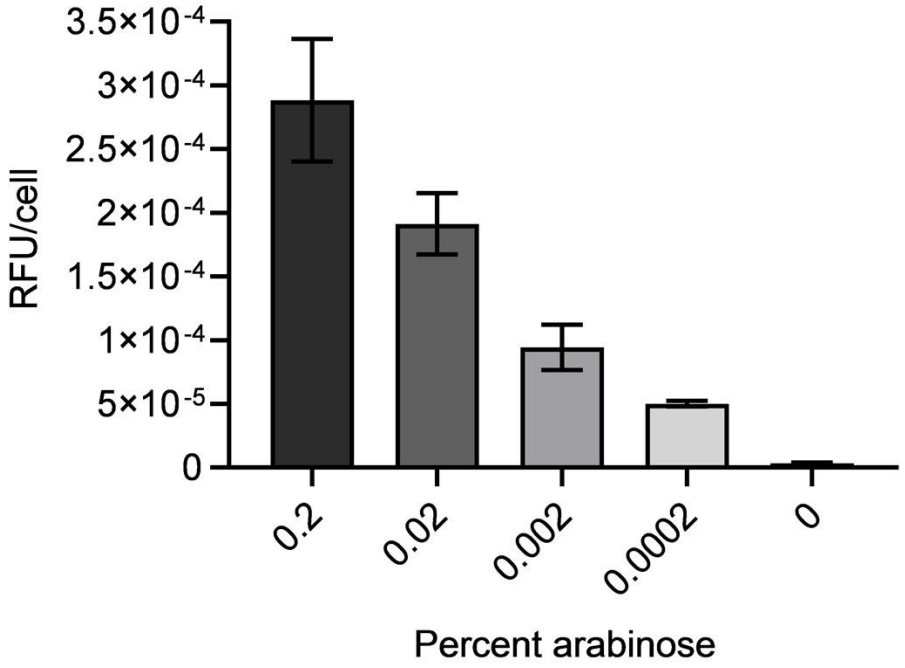

FIG 19 Modulation of GFP expression from the plasmid pGLO in *E.coli* LMG194. pGLO in *E. coli* LMG194 was treated with arabinose in four 10-fold dilutions or without arabinose. The relative fluorescence unit (RFU) per cell was calculated by dividing the total RFU by the number of cells (determined by the final OD_{595}) per arabinose treatment group. The fluorescence of the negative control (media alone) was subtracted from each value. Data are expressed as the mean $\pm$ SD of two biological replicates in quadruplicate. Statistical comparisons were made among all treatment groups and all samples were significantly different ($P \leq 0.05$) from each other using one-way ANOVA.

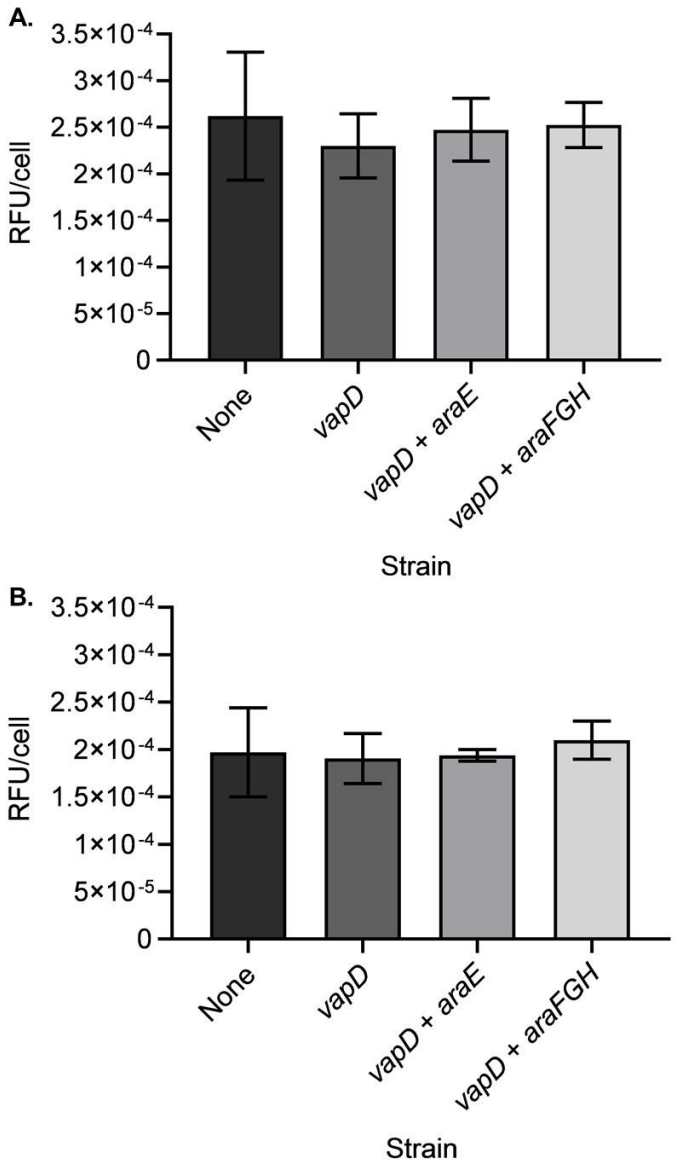

FIG 20 Transport of arabinose in NTHi. The amount of arabinose left in conditioned media was measured after growing NTHi strains with and without an additional arabinose permease to (A) mid-log phase (approximate OD_{600} of 0.4) or (B) stationary phase (approximate OD_{600} of 1.3) to measure if arabinose transport was affected by metabolism and/or cell growth. The cells were pelleted and the medium passed through a 0.2 micron filter. The control was media treated identically to the conditioned media, but without NTHi. The conditioned media and the control were used to induce pGLO in *E. coli* LMG194 to compare the amount of arabinose remaining in the conditioned media to the control media. RFU per cell was calculated by dividing the total RFU by the number of cells (determined by the final OD_{595}) for each strain. The fluorescence of the negative control (media alone) was subtracted from each value. Data are expressed as the mean $\pm$ SD of two biological replicates in quadruplicate of each strain of conditioned media. Statistical comparisons were made between the results of the conditioned media and the control and were not significantly different by one-way ANOVA.

DISCUSSION

The use of the P_{BAD} promoter with the regulatory functions of AraC is an efficient and highly useful way to study gene expression in a number of different bacterial strains. Expanding its use to organisms that live in niche environments, such as NTHi, provides another tool to study genes that aid bacteria in causing infections in human hosts and evading the immune response. A good example is the type II TA systems, which have been shown to be crucial for the survival of NTHi during infection, both *ex vivo* in primary human tissues and *in vivo* in the chinchilla model of acute otitis media (36, 60).

Other studies have successfully evaluated the use of the P_{trc} (146) and the P_{xylA} (149) in NTHi. The P_{trc} promoter is induced by IPTG, a synthetic molecule similar to allolactose, but is not hydrolyzed by β-galactosidase. Because IPTG is not degraded, concentrations remain constant during experiments. Use of the P_{xylA} promoter is a novel idea because, like the P_{BAD} promoter, its tight regulation is particularly important for the study of essential or toxic genes. However, this promoter is induced by xylose, a five-carbon sugar catabolized by many species, so it is important to verify catabolism genes in a species before utilizing this system. A number of studies have used the P_{xyl} from the particular species they are investigating and unlike promoters that are induced by IPTG and arabinose, there are no commercially-available plasmids that utilize P_{xyl} for gene expression.

In this study, *araC* with P_{BAD}::*vapD* was transformed into the chromosome of NTHi via homologous recombination at the site of a pseudogene. Two different arabinose permeases were also transformed into NTHi via homologous recombination at a bacteriophage terminase site to determine if an *E. coli* arabinose permease increased

arabinose transport across the membrane. Even when an arabinose permease was present, no significant amount of growth arrest upon induction of *vapD in cis* was observed (Fig. 18). It has been previously shown that the TA toxins VapC-1, VapD, and ToxA have ribonuclease activity (36, 46, 60, 85). It has also been shown that their corresponding TA operons contribute to survival of NTHi during infection (36, 60). However, it is important to note that the timeline of the assays utilized to measure effects on survival were over a period of 2-4 days. The first model, the EpiAirways (AIR 100 ABF; MatTek, Ashland, MA, USA), consists of primary human respiratory epithelial tissue at the air-liquid interface infected by NTHi over two (or more) days and is the closest *in vitro* model of the human upper airway that is currently available (108). In the second model, the chinchilla model of experimental otitis media, NTHi infection is quantified over four days and is extremely useful in the study of this and other auditory and respiratory tract infections due to the anatomical and physiological similarities between the chinchilla middle ear and that of humans (150). Through fluorescence measurements of pGLO in *E. coli* LMG194, the equal concentrations of arabinose in conditioned media compared to the arabinose-only control demonstrated that arabinose did not cross the NTHi membrane even with the addition of an *E. coli* arabinose permease (Fig. 20). It may be useful to explore *araC* with P_{BAD}::*vapD* and permease induction with arabinose on a plasmid *in trans* in the background of NTHi. However, it would difficult to induce the antitoxin, toxin, and a permease *in trans* because all three plasmids would need to be compatible in the background of NTHi. This is a challenging task due to the diversity of the NTHi genome (151). *H. influenzae* strains are naturally competent and can transport several hundred kilobases of foreign DNA fragments at a time from the environment (152, 153). The

advantages of DNA uptake include using the nucleotides as a food source, repair of damaged DNA, and recombination into the chromosome to acquire new genes that may be beneficial to cellular survival. The latter has implications in plasmid maintenance because new DNA that has homologously recombined into the chromosome of *H. influenzae* could be antibiotic resistance genes, new virulence factors, or origins of replication that may affect plasmid replication. A large challenge faced while conducting experiments with NTHi has been differences in plasmid survivability between two NTHi strains, namely 86-028NP and R2866. In fact, it was not possible to construct the *araC* with P_{BAD}::*vapD* recombinant at the targeted pseudogene site in 86-028NP. For reasons currently unknown, the 86-028NP strain generated only spontaneous mutants rather than recombinants (data not shown).

H. influenzae survives as a commensal in a niche environment, the upper respiratory tract of the human host, a unique setting. It can also act as a pathogen in the middle ear, nasopharynx, and lung (154). Finally, it can invade the bloodstream and the cerebrospinal fluid (CSF) as well (155). These areas differ greatly in nutrient and oxygen availability (156). For example, mucous membranes, biofilms, and the middle ear have limited oxygen levels (157) compared to epithelial surfaces (158). Serum-resistant invasive strains can disrupt epithelial cells and penetrate into the bloodstream and ultimately cause infection of the CSF. However, the CSF is 99% water, which limits nutrients for *H. influenzae* (159), and has lower oxygen tension compared to that of the bloodstream. Therefore, it is necessary for NTHi to maintain genes that help it quickly adapt to these changes in its microenvironment (160).

A study of metabolic diversity in *H. influenzae* determined that all species maintained enzymes for the Embden-Meyerhof-Parnas (EMP) pathway (with the exception of glucokinase), and all the enzymes in the pentose phosphate pathway for glucose degradation (PPP) (161). As a follow-on, researchers investigated the metabolic and physiological properties of four strains: one respiratory tract isolate (Hi2019), two invasive strains (R2866 and C188), and the reference strain (Rd KW20) (156, 161, 162). However, even though the central carbon metabolic pathway is highly conserved among strains, pathway preference and carbon source utilization differed (156, 161). While interesting, this result is not unexpected; NTHi must quickly adapt to changes in the host in order to survive. Researchers have shown that NTHi has transport capabilities for glucose, xylose, ribose, galactose, fructose, glycerol, fucose, pyruvate, and sialic acid (152, 163, 164). Although NTHi has the ability to transport and catabolize a range of carbon sources, no reports mention arabinose availability in areas NTHi colonizes. NTHi does not maintain homologues to the arabinose transport and catabolism genes; however, it may be possible for another pathway to transport arabinose. While the bioassay to determine arabinose transport provided evidence that R2866 did not transport arabinose (Fig. 20), further investigation into arabinose transport and catabolism in other strains of NTHi might be valuable, as the use of the arabinose-inducible P_{BAD} vectors would be extremely useful tools in the study of gene expression.

MATERIALS AND METHODS

Bacterial strains and culture conditions. The bacterial strains and plasmids used in these studies are listed in Table 6. *E. coli* strains were grown in LB broth or agar with or without 100 µg/mL ampicillin, 250 µg/ mL erythromycin, 20 µg/mL chloramphenicol, 100 µg/mL spectinomycin, or 10 µg/mL gentamicin, as required. *E. coli* strains were routinely cultured for 18 h at 37 °C. NTHi strains were grown in brain heart infusion broth or agar supplemented with 10 µg/mL heme-histidine and 10 µg/mL β-NAD (sBHI), chocolate agar with 5 U/mL bacitracin, or defined media (147) with or without glucose supplemented with 10 mM sodium pyruvate. NTHi were routinely cultured for 18 h at 37 °C with 5% CO_2. To construct the ectopic strains, recombinants were selected on chocolate agar plates with 2 µg/mL chloramphenicol, 25 µg/mL spectinomycin, or 5 µg/mL erythromycin. Single colonies were passaged for 3 days on selection plates before being confirmed by DNA sequencing of PCR-amplified purified genomic DNA.

Construction of an NTHi strain R2866 *ΔvapXD* mutant. The *vapXD* operon was previously deleted from NTHi strain 86-028NP (36). Due to the sequence homology between strains 86-028NP and R2866 flanking the *vapXD* operon (165), a PCR product from pDD788 (36) and the MIV transformation protocol (81) were used to create the R2866 Δ*vapXD* deletion mutant. Recombinants were selected on chocolate agar plates with chloramphenicol for three days. Genomic DNA preparations were used to characterize the deletion via PCR using primers that flanked the deletion.

To compare the growth dynamics, the parent and mutant strains were re-suspended from sBHI plates into fresh sBHI broth. The strains were grown to early log phase (OD_{600} of ~0.2 - 0.3) before being diluted into a 1:1 (vol:vol) ratio to a final volume

TABLE 5 Bacteria, plasmids and primers used in this study

Strains	**Description**	**Source**
DH5α	F– Φ80*lac*ZΔM15 Δ(*lac*ZYA-*arg*F) U169 *rec*A1 *end*A1 *hsd*R17 (rK–, mK+) *pho*A *sup*E44 λ– *thi*-1 *gyr*A96 *rel*A1	Laboratory collection
LMG194	F$^-$ Δ*lacX*74 *galE thi rpsL* Δ*phoA* (*PvuII*) Δ*ara*714 *leu*::Tn10	Invitrogen
DD12	MC4100 *recA::RP4-2-Tc::Mu*, kanamycin resistance	(146)
R2866	Blood isolate from an immunocompetent child with meningitis immunized with the Hib vaccine	A.L. Smith
R2866 Δ*vapXD*	*vapXD* TA locus replaced with a chloramphenicol resistance cassette	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD*	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD in cis*	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araE}::*araE*	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD* and P_{araE}::*araE in cis*	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araFGH}::*araFGH*	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD* and P_{araE}::*araFGH in cis*	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* + (pDD1233)	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD in cis* and *lacI*^{q} with P_{trc}*::vapX in trans*	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araE}::*araE* + (pDD1233)	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD* and P_{araE}::*araE in cis* and *lacI*^{q} with P_{trc}*::vapX in trans*	This work
R2866 Δ*vapXD* RS09695::*araC* P_{BAD}*::vapD* RS04985::P_{araFGH}::*araFGH* + (pDD1233)	R2866 Δ*vapXD* with *araC* with P_{BAD}*::vapD* and P_{araE}::*araFGH in cis* and *lacI*^{q} with P_{trc}*::vapX in trans*	This work
Plasmids	**Description**	**Source**
pBAD33	Highly regulated expression vector, chloramphenicol resistant	(77)
pTrcHis2A	Expression vector with C-terminal polyhistidine tag, ampicillin resistant	Invitrogen
pBluescript SK(+)	Cloning vector, ampicillin resistant	Laboratory Collection

TABLE 5 Continued

Plasmids	**Description**	**Source**
pSPEC1	Source of spectinomycin resistance cassette	(166)
pEJ18	Source of erythromycin resistance cassette	(107)
pGLO	pBAD-GFPuv, ampicillin resistant	Bio-Rad
pDD514	RSF1010 *ori* broad host range vector, gentamicin resistant	(146)
pDD788	pBluescript SK(+) with the *vapXD deletion*	(36)
pDD1020	5’ end of R2866_RS09695 in pBluescript SK(+)	This work
pDD1021	3’ end R2866_RS09695 in pDD1020	This work
pDD1022	R2866_RS09695 spectinomycin resistant -targeted delivery vector	This work
pDD1077	*vapX* in pTrcHis2A	This work
pDD1218	*vapD* in pBAD33	This work
pDD1227	pDD1022 with *araC* with P_{BAD}::*vapD*	This work
pDD1228	5’ end of R2866_RS04985 in pBluescript SK(+)	This work
pDD1229	3’ end R2866_RS04985 in pDD1229	This work
pDD1230	NTHI_RS04985 erythromycin resistant - targeted delivery vector	This work
pDD1233	*lacI*^{q} with P_{trc}::*vapX* in pDD514	This work
pDD1235	pDD1230 with *araE*	This work
pDD1236	pDD1230 with *araFGH*	This work
Primers	**Description**	**Source**
PvapDXba+SDFor	AAAATCTAGAATAAGGAAAATTATGTACG	Eurofins
pBAD33VapDHindRev	TAGGAAGCTTTTAATTTCGTATAAAATCAG	Eurofins
VapXSKSacFor	AAAAGAGCTCTATGGAACTAAGACAAC	Eurofins
86VapLexRev (*KpnI*)	ATAAGGTACCTTATTCAGCTTCGG	Eurofins
Xlacl for (X*baI*)	GTCATCTAGAAAACGCGCGAGGCAGC	IDT
Xhis rev (X*baI*)	TATTTGTCTAGAGGCAGTTCCCTACTCTCG	IDT
1356-1SacFor	AAAAGAGCTCAACTTAAACAAGCTC	Eurofins
1356-1SpeRev	AATTACTAGTAGCCTCTTCAATAC	Eurofins

TABLE 5 Continued

Primers	**Description**	**Source**
1356-2XhoFor	TTTCCTCGAGTTTAGCGGTGAATGG	Eurofins
1356-2KpnRev	AAAAGGTACCACACCCAACCACTTC	Eurofins
SpecPstFor	ACAACTGCAGAGATCTATCGATTTTCG	Eurofins
SpecPstRev	AAGCCTGCAGAGATCTGTACATCG	Eurofins
pBAD33BglIIRev	TTATAGATCTTATTGTCTCATGAG	Eurofins
pBAD33BglIIFor	TATCAGATCTGCATAATGTGC	Eurofins
RS04985-1SacFor	ATTTGAGCTCAAATTCCGACCGCTC	Eurofins
RS04985-1XbaRev	AGTATCTAGACCAAGTTTGAGATGC	Eurofins
RS04985-2XhoFor	AAGGCTCGAGTTGATGTGGCAGATG	Eurofins
RS04985-2KpnRev	AAAAGGTACCGGCGTAGCACATAAC	Eurofins
PermCSpeFor	AAATACTAGTAACACACACGCCATTCC	Eurofins
ErmCSpeRev	GTTAACTAGTGCAGTTATGCATCC	Eurofins
ParaESpeFor	GATGACTAGTATAGTGTTCCTTGCC	Eurofins
AraESpeStopRev	TTAAACTAGTTCAGACGCCGATATTT	Eurofins
ParaFGHSpeFor	AGGCACTAGTTATGGATTAATCTGC	Eurofins
AraFGHSpeRev	AAAAACTAGTTCAGACAGTGCGTTTCG	Eurofins

of 150 microliters of media in a 96-well Falcon polystyrene microplate (Corning, Corning, NY). The plate was covered with a gas permeable membrane, loaded into a Multiskan FC incubating plate reader (Thermo Fisher Scientific, Waltham, MA USA), and incubated at 35°C with a shaking protocol of five seconds on, five seconds off with measurements taken every 10 mins at OD_{595} for 7 h. Each growth assay included three biological replicates with three technical replicates.

Construction of artificially-inducible NTHi genes in *E. coli.* To transform this system into NTHi, the toxin, *vapD*, was cloned into plasmid pBAD33 so that it was under control of an arabinose-inducible promoter. NTHi strain 86-028NP genomic DNA was used to amplify *vapD* with PvapDXba+SDFor and pBAD33VapDHindRev primers. After digestion with *XbaI* and *HindIII*, the PCR fragment was ligated to pBAD33, resulting in pDD1218 (*vapD* under the control of the P_{BAD} promoter). VapXSKSacFor and 86VapLexRev primers were used to amplify *vapX* from 86-028NP genomic DNA. The PCR fragment was digested with *SacI* and *KpnI* and ligated to pTrcHis2A, resulting in pDD1077 (*vapX* under the control of the P_{trc} promoter). XlacIFor and XhisRev primers were used to amplify pDD1077. This PCR product was digested with *XbaI* and ligated to the broad-host range conjugal vector, pDD514 (146), resulting in pDD1233.

Construction of the *in cis* delivery systems and reconstituted NTHi strains. Based on previously published ectopic delivery vectors in NTHi (106, 143), two locations in the genome of R2866 were targeted for gene delivery. The *metH* pseudogene (R2866_RS09695) and the large subunit of the PBSX phage terminase gene (R2866_RS04985) were chosen as sites for gene delivery *in cis.* The large subunit of the PBSX phage terminase gene was chosen because R2866 only maintains a portion of the

PBSX bacteriophage genome. After scrutiny of the R2866 genome, it was determined that this gene would provide enough flanking regions to enable the homologous recombination of a large PCR product into the genome without impeding vital genes necessary for bacterial metabolism and replication.

The R2866_RS09695 targeted delivery vector was constructed by amplifying 899 bp of the 5' end of R2866_RS09695 from R2866 genomic DNA with 1356-1SacFor and 1356-1SpeRev primers and ligating the product into pBluescript SK(+) digested with *SacI* and *XbaI* (produces a compatible overhang which can be ligated to *SpeI*), resulting in pDD1020. The second homologous region of 953 bp of the 3' end of R2866_RS09695 was amplified from R2866 genomic DNA with 1356-2XhoFor and 1356-2KpnRev primers and ligated into pDD1020 digested with *XhoI* and *KpnI*, resulting in pDD1021. The last step was amplifying a spectinomycin resistance cassette from pSPEC1 (166) with SpecPstFor and SpecPstRev primers and ligating to pDD1021 digested with *PstI*, resulting in the delivery vector, pDD1022 (Fig. 16A). The primers pBAD33BglIIFor and pBAD33BglIIRev were used to amplify pDD1218. The PCR product was digested with *BglII* and ligated into pDD1022 digested with *BamHI* (produces a compatible overhang which can be ligated to *BglII*), resulting in pDD1227 (Fig. 16B). This arabinose-inducible targeted delivery vector resulted in the fusion of *araC* to the P_{BAD} promoter controlling *vapD*.

R2866 Δ*vapXD* was made chemically competent (81) and the R2866 Δ*vapXD* RS09695::*araC* with P_{BAD}*::vapD* mutant was constructed by homologous recombination of the PCR product using pDD1227 template. After selection on chocolate agar with 25 µg/mL spectinomycin, a single colony of each transformant was passaged three times on

antibiotic-containing plates. DNA sequencing of PCR-amplified genomic DNA preparations confirmed the correct strain.

The R2866_RS04985 delivery vector was constructed by amplifying 697 bp of the 5’ end of R2866_RS04985 from R2866 genomic with RS04985-1SacFor and RS04985-1XbaRev primers and ligating the product into pBluescript SK(+) digested with *SacI* and *XbaI*, resulting in pDD1228. The second homologous region of 592 bp of the 3’ end of R2866_RS04985 was amplified from R2866 genomic DNA with RS04985-2XhoFor and RS04985-2KpnRev primers and ligated into pDD1228 digested with *XhoI* and *KpnI*, resulting in pDD1229. The last step was amplifying an erythromycin resistance cassette from pEJ18 (107) with PermCSpeFor and ErmCSpeRev primers and ligating into pDD1229 digested with *SpeI*, resulting in the second delivery vector, pDD1230 (Fig. 17A).

Two different arabinose permeases were amplified from *E. coli* strain DH5α genomic DNA to individually clone into pDD1230. The first permease, the low-affinity arabinose transporter, *araE*, was amplified with ParaESpeFor and AraESpeStopRev primers. The second permease, the high-affinity arabinose transporter, *araFGH*, was amplified with ParaFGHSpeFor and AraFGHSpeRev primers. Both products were digested with *SpeI* and ligated into pDD1230, resulting in pDD1235 and pDD1236, respectfully (Fig. 17B and C).

R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD* was made chemically competent (81) to independently transform pDD1235 and pDD1236 into the chromosome, resulting in R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD* RS04985::P_{araE}::*araE* and R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD* RS04985::P_{araFGH}::*araFGH*. After selection on chocolate agar with 5 µg/mL erythromycin,

a single colony of each transformant was passaged three times on antibiotic-containing plates. PCR-characterized genomic DNA preparations confirmed each correct strain.

Construction of the *in trans* delivery systems into reconstituted NTHi strains. The antitoxin, VapX, was cloned into pDD1233 as previously mentioned. To induce VapD and VapX independently in the same strain, it was necessary to utilize two different promoters with different inducers. Due to the small, finely tuned genome size of many NTHi strains, it is particularly difficult to find unnecessary areas of the genome to utilize for targeted delivery systems. For this reason, the antitoxin in pDD1233 was conjugated into the newly constructed strains R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD*, R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD* RS04985::P_{araE}::*araE*, and R2866 Δ*vapXD* R2866_RS09695::*araC* with P_{BAD}::*vapD* RS04985::P_{araFGH}::*araFGH* (146). After selection on chocolate agar with 5 µg/mL gentamicin, transconjugates were passaged three times on antibiotic-containing plates. PCR-characterized plasmid preparations confirmed that pDD1233 was successfully transferred to each strain.

Modulation of expression of GFP in glucose-free defined media. The plasmid pGLO in *E. coli* LMG194 was used to measure the expression of GFP. The LMG194 strain with pGLO was re-suspended from glucose-free defined media with 10 mM sodium pyruvate agar and 100 µg/mL ampicillin into identical broth media without agar to an approximate OD_{600} of 0.16. The re-suspended culture was diluted into a 1:1 (vol:vol) ratio to a final volume of 150 microliters of medium in a 96-well Falcon polystyrene microplate (Corning, Corning, NY) containing identical media with 100 µg/mL ampicillin and 26.4 mM (0.4% wt/vol) arabinose in four 10-fold dilutions down to 0.0264 mM (0.0004% wt/vol) or

media without arabinose. The final induction concentrations were 13.2 (0.2% wt/vol), 1.32 mM (0.02% wt/vol), 0.132 mM (0.002% wt/vol), and 0.0132 mM (0.0002% wt/vol) arabinose. The plate was covered with a gas permeable membrane, loaded into a Multiskan FC incubating plate reader (Thermo Fisher Scientific, Waltham, MA USA), and incubated at 35 °C with a shaking protocol of five seconds on, five seconds off for 12 h. The final OD_{595} was recorded and florescence intensity was measured (405 nm excitation, 520 nm emission) on a FLUOstar OPTIMA Microplate Reader (BMG LABTECH, Ortenberg, Germany). For *E. coli*, each 0.1 OD unit of absorbance equals approximately 1.0×10^8 cells/mL (167). Using the final OD_{595}, the number of cells per treatment group were calculated to determine the RFU/cell. Each assay included two biological and four technical replicates.

Measuring the transport of arabinose in NTHi. The strains were re-suspended from glucose-free defined media with 10 mM sodium pyruvate agar plates into identical fresh media without agar at an initial OD_{600} of 0.1. They were immediately induced with 2.65 mM (0.04% wt/vol) arabinose and grown to mid-log phase (approximate OD_{600} of 0.4) or stationary phase (approximate OD_{600} of 1.3). When the final OD_{600} was reached, aliquots were centrifuged at 13,000 x g for 10 min to pellet cells and cell debris. The supernatant was filter-sterilized through a 0.2 micron polyethersulfone membrane. The control was treated identically to the conditioned media except it was not inoculated with NTHi. The sterilized supernatants (conditioned media and NTHi-free media control) were used in the pGLO in LMG194 fluorescence assays to determine if there was a difference in arabinose concentrations remaining in the conditioned media versus the control.

pGLO in LMG194 was re-suspended from glucose-free defined media with 10 mM sodium pyruvate agar and 100 µg/mL ampicillin into identical media without agar to approximately OD_{600} 0.16. The re-suspended culture was diluted into a 1:1 (vol:vol) ratio to a final volume of 150 microliters of media in a 96-well microplate with the conditioned media and control. The same protocol described in the pGLO fluorescence experiment was used to measure GFP fluorescence induced by the conditioned media and control. Each assay included three biological and four technical replicates.

Statistical analyses. Differences among multiple-group treatments were determined by the one-way or the repeated-measures ANOVA as appropriate using GraphPad Prism 8.1.2 (GraphPad Software, La Jolla, CA USA). A *P* value of ≤ 0.05 was considered statistically significant.

Acknowledgments. This work was supported in part by the National Institute on Deafness and other Communication Disorders cooperative agreement U01 DC014756 to DAD. The National Institute on Deafness and other Communication Disorders (NIDCD) had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results. This research is in accordance with protocols 17-007 and 15-006 approved by the Institutional Biosafety Committee (IBC), Old Dominion University, USA.

ANALYSIS OF THE GENOME ORGANIZATION OF FOUR TA LOCI IN 12 CLINICAL ISOLATES OF NONTYPEABLE *HAEMOPHILUS INFLUENZAE*

INTRODUCTION

In 1995, *H. influenzae* was the first free-living organism to have its genome sequenced (168). This sparked the “genome era” and since then, 204,683 partial and whole bacterial sequences, 735 being *Haemophilus influenzae* strains, have been deposited into Genbank (169). As the community shifts from Sanger sequencing to next-generation sequencing (NGS), it has become much more feasible for clinicians and researchers to quickly and accurately sequence the genomes of bacterial isolates in-house and analyze the data (170). These higher-throughput methods are useful for tracking epidemiological data and targeting current or potential outbreaks. NGS can also be a beneficial way to customize treatment plans by sequencing the causative agent(s) of major bacterial infections and tailoring antimicrobial therapy based on the resistance genes present in the isolates involved in the infection. Targeted therapy is predicted to become an increasing focus because unintended consequences of antibiotic use can alter the microbiome, activate toxin-antitoxin systems, establish of environmental reservoirs, and increase the speed in which bacteria develop resistance to current drugs (145). A list of FDA-approved antibiotics and their approval years was recently reported (145). There was an increase in approval of new antibiotics from the late-1930’s until the mid-1980’s (Fig. 21). After that time, there was a steady decline in FDA-approved antibiotics. The increase in antibiotic resistance and decreased support for antibiotic drug discovery has led researchers to investigate other antimicrobial therapy approaches. A potential target is the type II toxin-antitoxin (TA) locus.

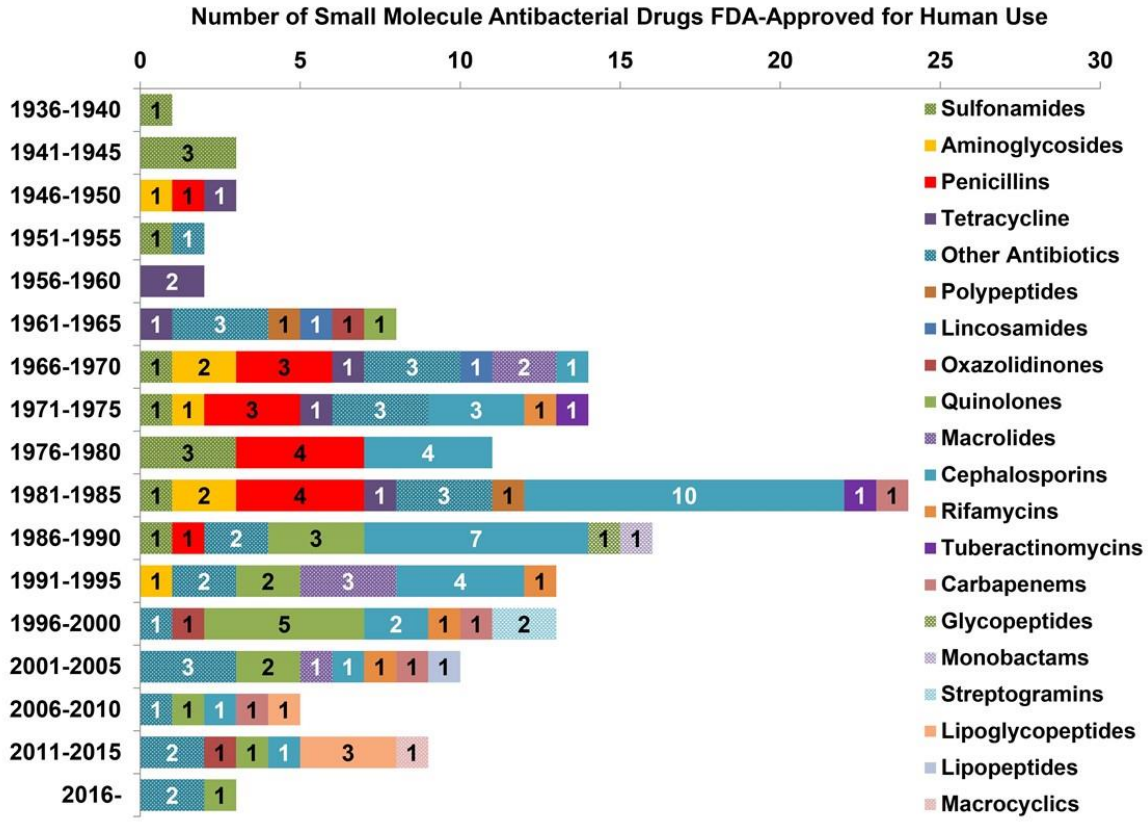

FIG 21 FDA approved antibacterial drugs for use in humans from 1936-2016. Antibacterial drugs approved by the US Food and Drug Administration (FDA) for use in humans. This plot shows the number of drugs approved within five-year periods from 1939 until 2017 (all drugs are listed in Table 1 of reference 145). The 157 drugs are grouped according to 20 classifications as shown by the legend. The greatest number of drugs were approved between 1981 and 1985. This image is reproduced from reference 145 reviewing FDA approved small molecule drugs, unintended consequences of antibiotic use, and the development and financial impact of antimicrobial therapy.

Nontypeable *Haemophilus influenzae* (NTHi) is a model organism to study one of the sub-families of the type II TA modules, the *vap* genes, because it maintains four *vap* operons in its very small (~2 Mbp) genome (36). This organism is human-adapted, meaning that there are no animal or environmental reservoirs for NTHi; it lives exclusively in humans. This trait has facilitated its evolution as a human commensal, living in the human upper airway without causing symptoms. Due to this lifestyle, NTHi strains have lost the genes necessary to synthesize many cofactors such as nicotinamide adenine dinucleotide (NAD) as well as heme, and scavenges these essential molecules from the host. Interestingly, however, this bacterium can also cause infections such as otitis media, bronchitis, and can invade the cerebrospinal fluid (CSF). These infection sites all vary in the amount of heme and NAD available.

NTHi is one of the leading causes of recurrent otitis media (17), infections that result in >$4 billion in health-care costs annually (18-20). This study sought to investigate the prevalence of type II TA modules in many NTHi strains as a potential therapeutic target. Type II TA loci contribute significantly to the survival and virulence of NTHi during infections and these loci increase bacterial persistence and dormancy following stress (36, 60). Due to the increase of antibiotic-resistant NTHi strains and the lack of a suitable vaccine to prevent these infections, the prevalence and genomic organization of four important, highly conserved type II TA loci in NTHi, *vapBC-1*, *vapBC-2*, *vapXD*, and *toxAvapA*, was investigated in 12 recently-sequenced clinical isolates. It was previously determined that a single mutation in the conserved region of the PIN domain in VapC-1 causes a significant decrease in survivability of NTHi strain 86-028NP (Fig.13) (143). The *vap* genes are part of pathogenicity islands (PAI), that are transferred between organisms

via horizontal gene transfer (HGT) to help them adapt to new ecological niches (171, 172). Continuing to study TA systems in *H. influenzae* is a novel way to exploit a potential antibacterial target. Due to the conservation of type II TA operons, this approach could be efficacious in the treatment of numerous pathogenic organisms.

RESULTS

Sequences. Identification and genomic analysis of *H. influenzae* has evolved from 16S rRNA genes sequencing, DNA-DNA hybridization and multilocus sequence typing (MLST), to whole-genome sequencing (WGS) facilitated by next-generation sequencing (NGS) (4, 22, 173-175). In an effort to contribute to genomic studies that have helped identify genes associated with virulence, growth, persistence, and nutrient acquisition, 12 clinical isolates of NTHi were sequenced (Table 6) on the Illumina™ MiSeq platform. Sequence assembly was completed in the Velvet de novo Assembly BaseSpace Labs application (Illumina) against NTHi strain 86-028NP (NC_007146.2). Strain 86-028NP (176) was chosen as the reference strain instead of Rd KW20 (168) because Rd is a laboratory strain with reduced virulence compared to the clinical isolate 86-028NP, and because 86-028NP has been validated in the chinchilla model of otitis media (177) and the primary human respiratory tissue model (EpiAirway, AIR 100 ABF; MatTek, Ashland, MA) (178).

Protein conservation and alignment. The genomes of the 12 recently-sequenced clinical isolates were used to investigate the conservation of type II TA modules in NTHi. Utilizing the PATRIC Basic Local Alignment Search Tool (BLAST) service (179), the 12 genomes were queried against the amino acid sequence of the antitoxin and toxin protein products of the following operons from the reference strain 86-028NP: *vapBC-1, vapBC-2, vapXD,* and *toxAvapA* (Table 8). The assumption was made that if the strain had both proteins in an operon, then it maintained the corresponding TA locus. After determining which isolates maintained the genes that produced the protein

products listed in Table 8, multiple sequence alignments were performed using Clustal Omega at EMBL-EBI (180-182) (Fig. 22A to H).

TABLE 6 Sequencing, assembly, and annotation metrics of NTHi strains sequenced on the MiSeq at Old Dominion University

Strain	Read Length (L)	No. of aligned reads (N)[a]	Coverage (x)[b]	Genome size (bp)	No. of contigs	Contig N50 (bp)	G+C content (%)	No. of CDS[c]
C294	151	1,053,389	83	1,807,236	79	92251	38.1	1842
C37	151	932,918	74	1,950,111	150	36486	38.1	2038
C447	151	965,862	76	1,916,993	92	59454	38.0	1951
C464	151	1,030,333	81	1,861,829	142	60580	38.0	1916
C470	151	940,232	74	1,894,592	97	57870	38.0	1922
R3001	151	1,025,145	81	1,826,634	89	72043	37.9	1830
R3168	151	590,531	47	1,958,029	144	91827	38.2	2002
R3171	151	552,672	44	1,780,770	83	101141	37.9	1761
R3172	151	501,645	40	1,810,932	72	79083	38.0	1791
R3175	151	536,811	42	1,788,908	102	74349	37.9	1781
R3176	151	532,429	42	1,908,985	90	114115	38.0	1955
R3177	151	456,581	36	1,921,437	103	67411	38.2	1988

[a]N = aligned reads for Read 1 and Read 2 combined
[b]Coverage = (L X N)/Reference genome bp (1,914,490 bp).
[c]Coding sequences

TABLE 7 Percent Identity Protein BLAST results against NTHi strain 86-028NP

	Protein							
Strain	**VapB-1**	**VapC-1**	**VapB-2**	**VapC-2**	**VapX**	**VapD**	**ToxA**	**VapA**
C294	X[a]	X[a]	100	100	97	93	100	100
C37	99	97	100	100	X[a]	X[a]	100	100
C447	X[a]	X[a]	100	99	97	99	100	100
C464	100	100	100	100	100	100	99	100
C470	100	100	100	99	98	47[b]	100	100
R3001	99	99	100	100	100	100	100	100
R3168	X[a]	X[a]	100	99	98	97	100	99
R3171	99	99	100	100	98	49[b]	100	100
R3172	99	99	100	99	98	47[b]	100	100
R3175	99	99	100	100	98	49[b]	99	88
R3176	100	100	100	99	100	100	100	100
R3177	99	99	X[a]	X[a]	98	99	98	94

[a]Strain is missing the TA locus
[b]Truncation in VapD

vapBC-1. Nine out of the 12 strains maintained the *vapBC-1* locus. Three of the strains (C464, C470, and R3176) shared 100 percent protein identity to the VapB-1 antitoxin (WP_005649046.1) and the VapC-1 toxin (WP_005649049.1) from the reference strain 86-028NP. The remaining six strains (C37, R3001, R3171, R3172, R3175, and R3177) were 99 percent identical to the 86-028NP VapB-1 antitoxin. They shared the same, semi-conservative amino acid change at position 3 (threonine to alanine). Five strains (R3001, R3171, R3172, R3175, and R3177) had a 99 percent identity score, and C37 had a 97 percent identity to the 86-028NP VapC-1 toxin. All six strains shared the same conserved mutation at position 21 (isoleucine to valine) and at position 36 (methionine to isoleucine). In strain C37, an additional amino acid mutation in VapC-1 is a conserved mutation at position 93 (arginine to glutamine).

vapBC-2. Eleven of the 12 strains maintain the *vapBC-2* locus; strain R3177 is the only one that does not. The 11 strains shared 100 percent identity to the 86-028NP VapB-2 antitoxin (WP_005648011.1). Strains C447, C470, R3168, R3172, and R3176 are 99 percent identical to the 86-028NP VapC-2 toxin (WP_005651560.1) and share the same non-conservative mutation at position 59 (valine to glycine), with the exception of strain R3168 which has a semi-conserved mutation at position 17 (arginine to histidine).

vapXD. Eleven of the 12 strains maintain the *vapXD* locus; strain C37 is the only one that does not. Three strains, C464, R3001, and R3176 are 100 percent identical to the 86-028NP VapX antitoxin (WP_011272059.1) and the VapD toxin (WP_011272058.1). Strains C470, R3168, R3171, R3172, R3175, and R3177 are 98 percent identical to 86-028NP VapX and strains C294 and C447 are 97 percent identical. All eight strains have a conserved mutation at position 55 (tyrosine to

histidine) in VapX and the two strains that are 97 percent identical (C294 and C447) have a semi-conserved mutation at position 36 (valine to alanine). Surprisingly, the VapD toxin sequence had the most variation. Strains C447 and R3177 are 99 percent identical and have a semi-conserved mutation at position 23 (glutamic acid to glutamine). Strain R3168 is 97 percent identical and has the same semi-conserved mutation at position 23 (glutamic acid to glutamine) as well as a non-conserved mutation at position 30 (alanine to glutamic acid), and a semi-conserved mutation at position 31 (valine to threonine). Strain C294 is 93 percent identical and has the semi-conserved mutation at position 23 (glutamic acid to glutamine), two deletions at positions 40 and 41 (threonine and glutamine), a non-conserved mutation at position 48 (methionine to threonine), a conserved mutation at position 61 (alanine to serine), and a conserved mutation at position 92 (asparagine to serine). Interestingly, four out of the 12 strains (C470, R3171, R3172, and R3175) have a 47 amino acid deletion from position 30 to 77 that was previously identified in five other strains (85). In addition to the truncation, strains C470 and R3172 also have the semi-conserved mutation at position 23 (glutamic acid to glutamine) mutation. Strains R3171 and R3175 have a non-conserved mutation at position 29 (glycine to arginine).

toxAvapA. This type II TA locus is interesting because unlike the previously described TAs, *toxAvapA* (a homologue of the host inhibition of growth locus *higBA* (183)) is arranged with the toxin directly upstream of the antitoxin. This gene pair exhibits two promoters, one for the toxin and a second for the antitoxin, a characteristic that likely contributes to the stability of the antitoxin. All 12 strains maintain the *toxAvapA* locus. Eight strains (C294, C37, C447, C470, R3001, R3171, R3172, and R3176) are 100

percent identical to both the 86-028NP ToxA toxin (WP_005650215.1) and the 86-028NP VapA antitoxin (WP_005650217.1). Strain C464 is 99 percent identical to ToxA with a non-conserved mutation at position 59 (leucine to serine). Strain R3175 is also 99 percent identical with a non-conserved mutation at position 79 (arginine to cysteine). Strain R3177 is 98 percent identical with a semi-conserved mutation at position 23 (threonine to alanine) and a non-conserved mutation at position 32 (tyrosine to serine). Strain 3168 is 99 percent identical to VapA with a semi-conserved mutation at position 39 (alanine to valine). Strain R3177 is 94 percent identical to VapA and has 6 mutations: conserved at position 26 (serine to asparagine); semi-conserved at position 60 (valine to alanine); conserved at position 79 (aspartic acid to asparagine); non-conserved at position 95 (proline to serine); non-conserved at position 101 (aspartic acid to glycine); and semi-conserved at position 104 (threonine to alanine). Interestingly, R3177 has a 13 amino acid truncation at N-terminus of VapA, however, ToxA is maintained at its full length. A BLAST search based upon the protein sequence does not reveal any more strains with this 13 amino acid truncation (179).

A.

```
86-028NP    MLTKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
C37         MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
C464        MLTKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
C470        MLTKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3001       MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3171       MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3172       MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3175       MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3176       MLTKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
R3177       MLAKVFQSGNSQAVRIPMDFRFDVDTVEIFRKENGDVVLRPVSKKTDDFLALFEGFDETF    60
            **:************************************************************

86-028NP    IQALEARDDLPPQERENL    78
C37         IQALEARDDLPPQERENL    78
C464        IQALEARDDLPPQERENL    78
C470        IQALEARDDLPPQERENL    78
R3001       IQALEARDDLPPQERENL    78
R3171       IQALEARDDLPPQERENL    78
R3172       IQALEARDDLPPQERENL    78
R3175       IQALEARDDLPPQERENL    78
R3176       IQALEARDDLPPQERENL    78
R3177       IQALEARDDLPPQERENL    78
            ******************
```

FIG 22 VapBC-1, VapBC-2, VapXD, and ToxAVapA CLUSTAL O (1.2.4) multiple sequence alignments (MSA). The protein sequences belonging to four virulence associated protein (*vap)* operons from NTHi strain 86-028NP were used to BLAST 12 clinical isolates recently sequenced at Old Dominion University. Strains are listed in the following order for all MSAs: 86-028NP, C294, C37, C447, C464, C470, R3001, R3168, R3171, R3172, R3175, R3176, R3177. Strains not included in the alignment are marked with a caret in the figure legend. MSAs of the following proteins VapB-1 C294^, C447^, and R3168^ (A); VapC-1 C294^, C447^, and R3168^ (B); VapB-2 R3177^ (C); VapC-2 R3177^ (D); VapX C37^ (E); VapD C37^ (F); ToxA (G); VapA (H). An asterisk indicates positions which have a single, fully conserved residue; a colon indicates conservation between groups of strongly similar properties; a period indicates conservation between groups of weakly similar properties according to the Gonnet PAM 250 matrix. The colors represent physiological properties of the amino acids: red (small, hydrophobic), blue (acidic), magenta (basic), green (hydroxyl, sulfhydryl, amine), grey (unusual amino/imino).

B.

```
86-028NP        MIYMLDTNIIIYLMKNRPKIIAERVSQLLPNDRLVMSFITYAELIKGAFGSQNYEQSIRA      60
C37             MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
C464            MIYMLDTNIIIYLMKNRPKIIAERVSQLLPNDRLVMSFITYAELIKGAFGSQNYEQSIRA      60
C470            MIYMLDTNIIIYLMKNRPKIIAERVSQLLPNDRLVMSFITYAELIKGAFGSQNYEQSIRA      60
R3001           MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
R3171           MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
R3172           MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
R3175           MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
R3176           MIYMLDTNIIIYLMKNRPKIIAERVSQLLPNDRLVMSFITYAELIKGAFGSQNYEQSIRA      60
R3177           MIYMLDTNIIIYLMKNRPKIVAERVSQLLPNDRLVISFITYAELIKGAFGSQNYEQSIRA      60
                ************************:**************:************************

86-028NP        IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
C37             IELLTERVNVLYPNEQICLHYGKWANTLKKQGQPIGNNDLWIACHALSLNAVLITHNVKE      120
C464            IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
C470            IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3001           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3171           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3172           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3175           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3176           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
R3177           IELLTERVNVLYPNEQICLHYGKWANTLKKQGRPIGNNDLWIACHALSLNAVLITHNVKE      120
                ********************************:***************************

86-028NP        FQRITDLQWQDWTK    134
C37             FHRITDLQWQDWTK    134
C464            FQRITDLQWQDWTK    134
C470            FQRITDLQWQDWTK    134
R3001           FQRITDLQWQDWTK    134
R3171           FQRITDLQWQDWTK    134
R3172           FQRITDLQWQDWTK    134
R3175           FQRITDLQWQDWTK    134
R3176           FQRITDLQWQDWTK    134
R3177           FQRITDLQWQDWTK    134
                *:************
```

FIG 22 Continued.

C.

```
86-028NP        MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
C294            MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
C37             MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
C447            MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
C464            MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
C470            MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3001           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3168           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3171           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3172           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3175           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
R3176           MIEASVFMTNRSQAVRLPAEVRFSEEIKKLSVRVSGSDRILSPLNQSWDSFFLNDQAVSD   60
                ************************************************************

86-028NP        DFMNEREIAFQPEREAL 77
C294            DFMNEREIAFQPEREAL 77
C37             DFMNEREIAFQPEREAL 77
C447            DFMNEREIAFQPEREAL 77
C464            DFMNEREIAFQPEREAL 77
C470            DFMNEREIAFQPEREAL 77
R3001           DFMNEREIAFQPEREAL 77
R3168           DFMNEREIAFQPEREAL 77
R3171           DFMNEREIAFQPEREAL 77
R3172           DFMNEREIAFQPEREAL 77
R3175           DFMNEREIAFQPEREAL 77
R3176           DFMNEREIAFQPEREAL 77
                *****************
```

FIG 22 Continued.

D.

```
86-028NP      MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
C294          MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
C37           MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
C447          MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAGI   60
C464          MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
C470          MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAGI   60
R3001         MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
R3168         MLKYMLDTNIVIYVIKHRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
R3171         MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
R3172         MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAGI   60
R3175         MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAVI   60
R3176         MLKYMLDTNIVIYVIKRRPLEILSRFNQNAGKMCVSSITVAELYYGAEKSEYPERNIAGI   60
              ****************:****************************************** *

86-028NP      EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
C294          EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
C37           EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
C447          EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
C464          EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
C470          EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3001         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3168         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3171         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3172         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3175         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
R3176         EDFLSRLTILDYQPKHAAHFGNIKAELSKQGKLIGENDIHIAAHARSEGLILVSNNLREF   120
              ************************************************************

86-028NP      ERVIALRTENWV   132
C294          ERVIALRTENWV   132
C37           ERVIALRTENWV   132
C447          ERVIALRTENWV   132
C464          ERVIALRTENWV   132
C470          ERVIALRTENWV   132
R3001         ERVIALRTENWV   132
R3168         ERVIALRTENWV   132
R3171         ERVIALRTENWV   132
R3172         ERVIALRTENWV   132
R3175         ERVIALRTENWV   132
R3176         ERVIALRTENWV   132
              ************
```

FIG 22 Continued.

E.

```
86-028NP      MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSYLLEDP   60
C294          MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDALVNITLLQSGEKDIYRLSHLLEDP   60
C447          MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDALVNITLLQSGEKDIYRLSHLLEDP   60
C464          MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSYLLEDP   60
C470          MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
R3001         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSYLLEDP   60
R3168         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
R3171         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
R3172         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
R3175         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
R3176         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSYLLEDP   60
R3177         MELRQQIPTGCIKQFGQFGVPYVVGEVAEFLPDGDVLVNITLLQSGEKDIYRLSHLLEDP   60
              **********************************.******************:*****

86-028NP      EAE       63
C294          EAE       63
C447          EAE       63
C464          EAE       63
C470          EAE       63
R3001         EAE       63
R3168         EAE       63
R3171         EAE       63
R3172         EAE       63
R3175         EAE       63
R3176         EAE       63
R3177         EAE       63
              ***
```

F.

```
86-028NP      MYAIAFDLVVKDTQDYHPKGVQEAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
C294          MYAIAFDLVVKDTQDYHPKGVQQAYTDIGAVLAKFGFVR--GSLYTNTNEDMANLFQAMN   58
C447          MYAIAFDLVVKDTQDYHPKGVQQAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
C464          MYAIAFDLVVKDTQDYHPKGVQEAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
C470          MYAIAFDLVVKGTQDYHPKGVQQAYTDIG-----------------------------   29
R3001         MYAIAFDLVVKDTQDYHPKGVQEAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
R3168         MYAIAFDLVVKDTQDYHPKGVQQAYTDIGETLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
R3171         MYAIAFDLVVKDTQDYHPKGVQEAYTDIR-----------------------------   29
R3172         MYAIAFDLVVKDTQDYHPKGVQQAYIDIG-----------------------------   29
R3175         MYAIAFDLVVKDTQDYHPKGVQEAYTDIR-----------------------------   29
R3176         MYAIAFDLVVKDTQDYHPKGVQEAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
R3177         MYAIAFDLVVKDTQDYHPKGVQQAYTDIGAVLAKFGFVRTQGSLYTNMNEDMANLFQAMN   60
              **********.**********:** **

86-028NP      ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
C294          SLKQLAWISQSVRDIRAFRIEQWSDFTDFIRS  90
C447          ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
C464          ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
C470          ---------------AFRIEQWSDFTDFIRN  45
R3001         ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
R3168         ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
R3171         ---------------AFRIEQWSDFTDFIRN  45
R3172         ---------------AFRIEQWSDFTDFIRN  45
R3175         ---------------AFRIEQWSDFTDFIRN  45
R3176         ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
R3177         ALKQLAWISQSVRDIRAFRIEQWSDFTDFIRN  92
                       ***************.
```

FIG 22 Continued.

G.

```
86-028NP      MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
C294          MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
C37           MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
C447          MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
C464          MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHSE      60
C470          MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3001         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3168         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3171         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3172         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3175         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3176         MFNLKREHFRDDYLYRFYQYGDTHSKIPSNLYKVLARKLDMISASENINDLRSPPANHLE      60
R3177         MFNLKREHFRDDYLYRFYQYGDAHSKIPSNLSKVLARKLDMISASENINDLRSPPANHLE      60
              **********************:******** ************************** *

86-028NP      LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
C294          LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
C37           LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
C447          LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
C464          LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
C470          LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3001         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3168         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3171         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3172         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3175         LLEPKENKIYSIRVNKQYCLIFKYENNEVNNLYLDPHSYNL 101
R3176         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
R3177         LLEPKENKIYSIRVNKQYRLIFKYENNEVNNLYLDPHSYNL 101
              ****************** **********************
```

FIG 22 Continued.

H.

```
86-028NP      MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
C294          MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
C37           MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
C447          MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
C464          MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
C470          MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
R3001         MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
R3171         MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
R3172         MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
R3168         MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTVSNIVNNSSRITLEMAVKLAKV   60
R3175         ------------QEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   47
R3176         MMTRKPTSVGEILQEEFLEPLSLKISDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKV   60
R3177         MMTRKPTSVGEILQEEFLEPLSLKINDLAQILDVHRNTASNIVNNSSRITLEMAVKLAKA   60
                           ************.************.********************.

86-028NP      FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
C294          FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
C37           FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
C447          FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
C464          FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
C470          FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3001         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3171         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3172         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3168         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3175         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   94
R3176         FDTTPEFWLNLQTRIDLWDLEHNKRFQQSLANVKPALHRHDTSTFAM   107
R3177         FDTTPEFWLNLQTRIDLWNLEHNKRFQQSLANVKSALHRHGTSAFAM   107
              ******************:*************** *****.**:***
```

FIG 22 Continued.

TA loci organization. By using the protein BLAST service (179), it was determined that seven of the strains (C454, C470, R3001, R3171, R3172, R3175, and R3176) had all four *vap* operons and the remaining five strains (C294, C37, C447, R3168, and R3177) were missing one operon each. Strains C294, C447, and R3168 were missing *vapBC-1*; strain R3177 was missing *vapBC-2*; and strain 37 was missing *vapXD*. The next step was to determine if the four *vap* loci were located on the genome in the same region as the reference strain 86-028NP. Location was considered an important factor because even though NTHi is genetically diverse, the TA loci are important for bacterial survival and therefore conservation and maintenance are necessary. If the TA loci were not in the same regions as the reference genome, it could indicate that the genes were acquired at different times and by different means, such as by horizontal gene transfer from other microorganisms or from a bacteriophage. A region of one kilobase directly upstream and downstream of each *vap* operon was used in the PATRIC BLAST service (179) to query the 12 recently-sequenced strains to determine if the organization of the identified operons was the same as the reference sequence (Table 8). It was determined that all four *vap* operons shared the same flanking regions as in the reference strain 86-028NP.

Phylogenetic tree. Phylogenetic trees are important because they can illustrate evolutionary relationships among biological species. These ancestral relationships can offer some insight into potential treatments against infections caused by NTHi. For example, two recently diverged strains would have more nucleotides in common than two strains from different clades. While a phylogenetic tree will not differentiate what those similarities and differences are, exploring the genomes and past clinical records of the

TABLE 8 Operon location BLAST results against a two kilobase region around each *vap* operon from 86-028NP

	Gene			
Strain	*vapBC-1*	*vapBC-2*	*vapXD*	*toxAvapA*
C294	N/A[a]	Conserved	Conserved	Conserved
C37	Conserved	Conserved	Conserved	Conserved
C447	N/A[a]	Conserved	Conserved	Conserved
C464	Conserved	Conserved	Conserved	Conserved
C470	Conserved	Conserved	Conserved	Conserved
R3001	Conserved	Conserved	Conserved	Conserved
R3168	N/A[a]	Conserved	Conserved	Conserved
R3171	Conserved	Conserved	N/A[a]	Conserved
R3172	Conserved	Conserved	Conserved	Conserved
R3175	Conserved	Conserved	Conserved	Conserved
R3176	Conserved	Conserved	Conserved	Conserved
R3177	Conserved	N/A[a]	Conserved	Conserved

[a]N/A = not available

strains may provide insight on virulence factors.

The Codon Tree application in PATRIC (184) was used to build two phylogenetic trees, then upload them into the Interactive Tree of Life (ITOL) (185). The first tree was built with the 12 clinical isolates recently sequenced at ODU (Fig. 23). This tree was generated using 1000 protein sequences (a total of 399,105 aligned amino acids) and 953 gene sequences (1,130,649 aligned nucleotides). The second tree included 27 additional *H. influenzae* sequences collected from GenBank (169) (Table 9). This tree was generated using 977 protein sequences (a total of 309,629 aligned amino acids) and 900 gene sequences (845,211 aligned nucleotides). Two trees were built using PATRIC software to observe how the clinical isolates are distributed among a sampling of the *H. influenzae* population (Fig. 24). Many of the 27 typeable and nontypeable strains chosen have been previously studied, and a mixture of invasive and non-invasive strains was included to investigate the relationship of the recently-sequenced strains to the NTHi population (Fig. 24).

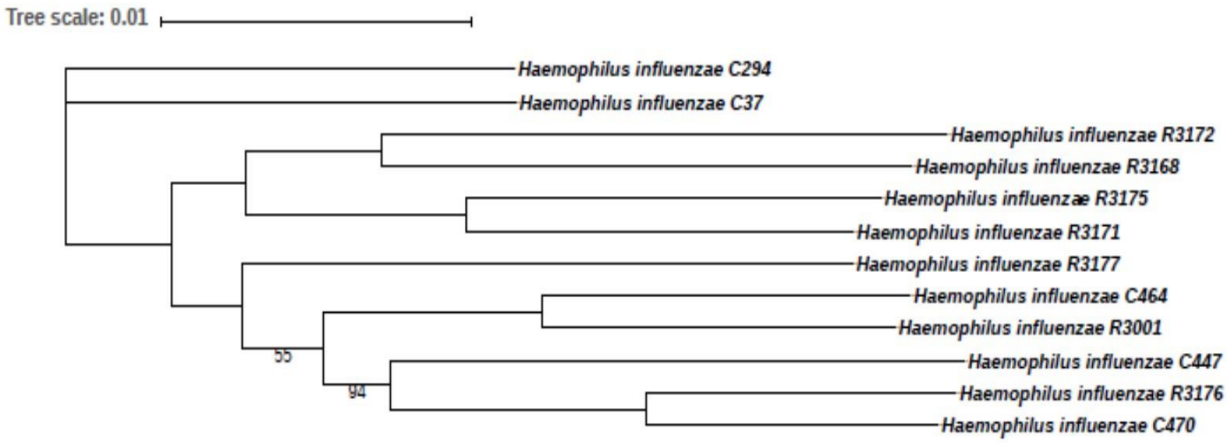

FIG 23 Phylogenetic tree of 12 NTHi clinical isolates recently sequenced at Old Dominion University. The Codon Tree application in PATRIC was used to build a phylogenetic tree with 12 clinical isolates of NTHi using 1000 protein sequences (a total of 399,105 aligned amino acids) and 953 gene sequences (1,130,649 aligned nucleotides). The tree was viewed in iTOL. Only bootstrap values of less than 100 are reported.

TABLE 9 Strains used to generate phylogenetic trees

Strain	Type	Description	Source
C294[a]	NT	Pericardium	A.L. Smith
C37[a]	NT	Sputum from a patient with chronic lung disease	A.L. Smith
C447[a]	NT	Ear	A.L. Smith
C464[a]	NT	Ear	A.L. Smith
C470[a]	NT	Ear	A.L. Smith
R3001[a]	NT	Bronchial lavage of a pediatric cystic fibrosis patient	A.L. Smith
R3168[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
R3171[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
R3172[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
R3175[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
R3176[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
R3177[a]	NT	Blood from Hib vaccinated patient	A.L. Smith
2019	NT	Sputum	(186)
60294N1	NT	Nasopharyngeal isolate	(187)
723	NT	OME isolate	(173)
86-028NP	NT	Nasopharyngeal isolate from a pediatric OM patient	R.S. Munson, Jr.
CGSHiCZ412602	NT	OME isolate	
NCTC 8143	NT	Type strain (ATCC 33391)	
PitAA	NT	OME isolate	(188)
PitEE	NT	OME isolate	(188)
PitGG	NT	OME isolate	(188)
PitHH	NT	OME isolate	(188)
PitII	NT	OME isolate	(188)
R2846	NT	AOM patient	A.L. Smith
R2866	NT	Blood isolate of a child with meningitis immunized the with Hib vaccine	A.L. Smith
R3021	NT	22.4-21	(21)
NML-Hia-1	a	Blood isolate	(189)
Hi609	a	Blood isolate	(190)
ATCC 10211	b	AMC 36-A-[572]	
10810	b	Isolated from a meningitis patient	(191)
NCTC 11394	b	Unknown	

TABLE 9 Continued

Strain	Type	Description	Source
NCTC 13377	b	Unknown	
M12125	c	Unknown	
Rd KW20	NT	Type d derivative	(168)
Hi467	e	CSF of meningitis patient	(192)
KR494	f	Necrotizing myositis	(193)
WAPHL1	f	Blood isolate	(194)
M21384	hybrid	Unknown	(195)
NCTC 11426	hybrid	Unknown	(195)

[a]Strains sequenced at Old Dominion University

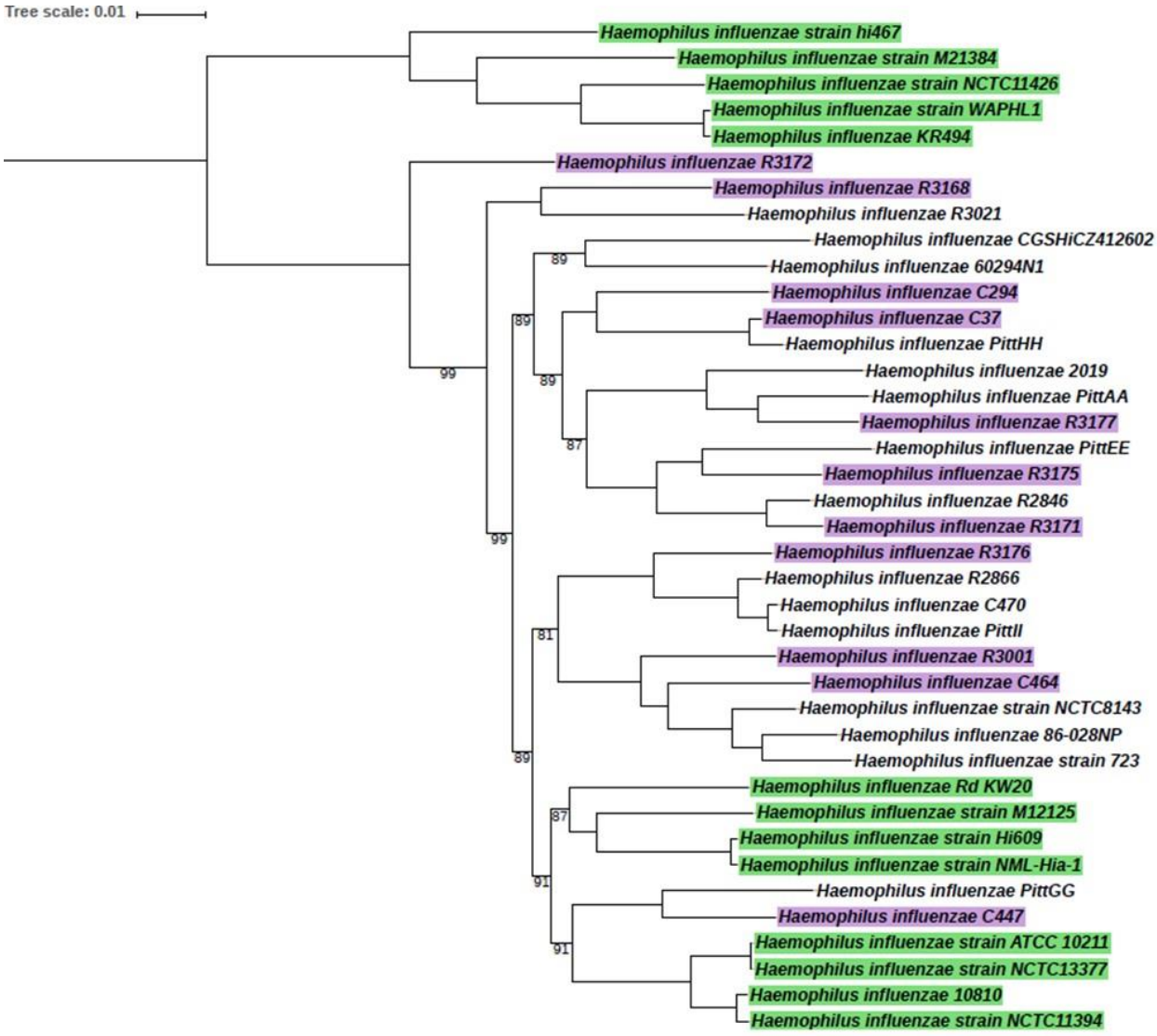

FIG 24 Phylogenetic tree of 39 strains of *H. influenzae*. Twelve NTHi clinical isolates recently sequenced at ODU and 27 typeable and nontypeable invasive and noninvasive sequenced strains collected from GenBank were used to generate a phylogenetic tree with the Codon Tree application in PATRIC using 977 protein sequences (a total of 309,629 aligned amino acids) and 900 gene sequences (845,211 aligned nucleotides). The tree was viewed in iTOL. Only bootstrap values of less than 100 are reported. Strains highlighted in purple are the 12 isolates recently sequenced at ODU. Strains highlighted in green are encapsulated strains from the 27 sequences collected from GenBank. The strains not highlighted are unencapsulated strains from the 27 sequences collected from GenBank.

DISCUSSION

H. influenzae is a host-adapted human pathogen, and genomic studies have determined that approximately two-thirds of its proteins have homologues in *E. coli*. Most of the genome reduction is due to the loss of paralogues (196). This evolution is advantageous to an obligate parasite of the human upper respiratory tract because it is metabolically expensive to maintain duplicate genes or genes that transport and catabolize products the organism does not need for survival and may never encounter in its native microenvironment. Encapsulated strains, especially *H. influenzae* serotype type b (Hib), are significantly more clonal than NTHi (173, 197-199). This could be the result of the protective nature of the capsule in typeable strains, which does not facilitate the acquisition of new DNA via horizontal gene transfer in the DNA-rich environment of their human hosts. However, NTHi strains are naturally competent, and their high rate of DNA uptake and homologous recombination significantly influences the genetic diversity of the species (175, 200). Twelve clinical isolates were sequenced to learn more about the genetic diversity and to investigate the conservation of four TA loci: *vapBC-1, vapBC-1, vapXD,* and *toxAvapA*.

Trillions of microorganisms cohabitate with their human hosts all over the body in various anatomical sites (gut, mouth, skin, vagina, upper respiratory tract) (201). Microbiome and normal flora are terms used to describe this harmonious relationship, and this niche occupation helps prevent the overgrowth of pathogens. However, stress imposed on the body, such as that caused by an infection, antibiotic use or a change in diet, can disrupt the normal flora (202). This disruption can result in both commensal and opportunistic organisms becoming pathogenic by either increasing in numbers or

migrating into normally sterile sites and causing infections (203). NTHi is an opportunistic organism, and a study by Kaur *et al.* determined no major phylogenetic differences between asymptomatic colonizing strains and strains that caused acute otitis media (AOM) (204). Studies have linked TAs to the survival of environmental and pathogenic bacterial species, which is why it is so important to continue to investigate the conservation of these operons and exploit them as potential antimicrobial drug targets (205, 206).

This study sequenced 12 invasive and noninvasive clinical isolates of NTHi from a range of infection sites in the body including: two sputum, three ear, and seven blood isolates. Using those sequences and various applications in PATRIC, it was determined that all 12 isolates had at least three of the four *vap* operons of interest, and seven of these strains (58%) maintained all four TAs. In the five strains that maintained three of the *vap* operons, *vapBC-1* was not maintained in three, *vapBC-2* was not maintained in one, and *vapXD* was not maintained in one. The percent identity of the antitoxin and toxin proteins compared to the sequences of strain 86-028NP was often 98-100 percent. Five strains maintained three of the four TAs of interest. Three of these strains are considered invasive because they infected normally sterile sites (C294, R3168, and R3177), and two of these strains were considered noninvasive commensals (C37 and C447). Interestingly, none of the mutations in VapC-1 were in the PilT N-terminal (PIN) domain (62, 65). The largest difference in the amino acid sequence was a 47 amino acid deletion in VapD in four of the strains. This unique in-frame deletion was discovered in ten other clinical strains during a study that identified and characterized *vapXD* in 59 commensal and disease-causing isolates NTHi from the nasopharynx,

sputum/tracheal aspirate/ear, and blood/CSF (85). Daines *et al.* (2004) determined that this deletion occurred in 7% of the nasopharynx, 50% of the sputum/tracheal aspirate/ear, and 29% of the blood/CSF isolates that they sampled. Interestingly, this deletion occurred in 20% of the sputum/tracheal aspirate/ear, and 43% of the blood isolates that were sampled. The second sputum isolate that was sequenced did not maintain the entire *vapXD* operon. The significance of this truncation is not yet known, however, future studies to determine its effect on survivability and virulence will allow investigators to gain insight into the unusual conservation of this in-frame deletion. It would be interesting to explore the implications of TA loci in invasive, serum-resistant NTHi strains. Normal human serum is bactericidal, but serum-resistant NTHi prevent the accumulation of the membrane attack complex (MAC), thus preventing the complement system from removing the pathogen. The novel system designed to study the *vapC-1* mutations *in cis* (Fig. 12 and 13) could be utilized to compare the survivability of NTHi expressing either the full length or the truncated VapD toxin gene in the EpiAirway or chinchilla models.

The genomic location of the TAs based on their flanking one-kilobase regions were identical, indicating that all 12 strains shared the genomic organization of at least those four regions, which may be important to the conservation of the operons. This information continues to support the notion of the necessity for microorganisms to maintain TAs to facilitate adaptation to changing environments. Future studies to expand the investigation into the organization of the TA loci with regards to a larger region in the genome would also be useful for determining any positive or negative effects on bacterial survival. While it has been determined that virulent epidemic

bacteria have more TA loci than less virulent bacteria (207), a majority of bacterial species have between one and five or six to ten type II TA loci (208). NTHi may have efficiently reduced its genome to coexist with the host as a commensal, but maintained the necessary genes to survive during times of stress. The importance of these TA loci to survival and virulence and their high level of conservation provide excellent support for the approach of targeting these operons to prevent infections caused by non-b serotypes, NTHi strains, and other pathogens that use this mechanism to survive during infection.

MATERIALS AND METHODS

Sequencing. Genomic DNA was extracted from all 12 strains using the Thermo Scientific GeneJet™ Genomic DNA Purification Kit (Waltham, MA, USA) according to the manufacturer's instructions. Sequencing libraries were prepared from genomic DNA using Nextera™ XT DNA Library Preparation Kit (Illumina, San Diego, CA, USA) and samples were run separately for sequencing with the MiSeq® 300-cycle micro kit (version 2) on an Illumina™ MiSeq platform. Trimming and *de novo* assembly were completed in the Velvet de novo Assembly BaseSpace Labs application (Illumina) against NTHi strain 86-028NP (NC_007146.2). Automatic annotation was done by Rapid Annotations using Subsystems Technology (RAST) v2.0 (209, 210) and the Pathosystems Resource Integration Center (PATRIC) (211) (Table 6). Read 1 and Read 2 were combined to determine the number of aligned reads. The coverage was calculated with the following equation (read length (L) x number of aligned reads (N))/reference sequence genome size (bp). The genome size of strain 86-028NP is 1,914,490 bp (169).

Phylogenetic inference. Twelve clinical isolates sequenced in-house plus an additional 27 sequenced typeable and nontypeable strain sequences collected from GenBank (169) were used to generate phylogenetic trees using the Codon Trees function in PATRIC (184) (Table 10). The Codon Trees function uses amino acid and nucleotide sequences from PATRIC's global Protein Families (PGFams) (184). The protein sequences were aligned using multiple sequence comparison by log-expectation (MUSCLE) (212), and the nucleotide sequences were aligned using the Codon_align function of BioPython (213). Randomized Axelerated Maximum Likelihood (RAxML) was used to analyze and describe the protein and nucleotide alignments (214) and rapid

bootstrap support values were generated with RAxML software (215). Newick files were uploaded to the Interactive Tree of Life (ITOL) (185) to view the final product.

Acknowledgments. This work was supported in part by the National Institute on Deafness and other Communication Disorders cooperative agreement U01 DC014756 to DAD. The National Institute on Deafness and other Communication Disorders (NIDCD) had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; and in the decision to publish the results. This research is in accordance with protocols 17-007 and 15-006 approved by the Institutional Biosafety Committee (IBC), Old Dominion University, USA.

CONCLUSION

These three studies provide insight into the importance of the type II TA loci in NTHi. In the first study, the crystal structure of VapBC-1 from NTHi was solved in a collaboration with investigators at the University of Kansas. Utilizing this information, aspartate-to-asparagine and glutamate-to-glutamine single and double mutations of four conserved residues in the PIN domain were constructed. Both the homo- and hetero-interactions of the mutant VapC-1 toxins in *E. coli* as well as the mutants' ability to cause growth arrest were quantified. All VapC-1 mutants interacted significantly with themselves (homodimerization) and with wild-type VapB-1 (heterodimerization). Only one mutant, VapC-1 D99N, caused significant growth arrest in the *E. coli* assay. The *vapC-1* mutants were expressed with wild-type *vapB-1* under the control of the native *vapBC-1* promoter in an infection model utilizing primary human epithelial tissues. It was determined that a single mutation to the conserved amino acids of the PIN domain attenuated NTHi survival in this *ex vivo* model. Not only were the mutant strains attenuated for survival, they also displayed survival numbers that were not significantly different from the 86-028NP Δ*vapBC-1* knockout that was previously reported (36). Another important finding was that additional mutations in VapC-1 were not additive. This result will be helpful in efforts to determine a drug target, because it suggests that only one small area of the domain needs to be targeted to interfere with ribonuclease activity. This study was the first to determine the crystal structure of VapBC-1 from NTHi as well as the first to utilize a model to study TA mutant operons in the background of NTHi during infections of primary human tissues *ex vivo*. Promising future steps for this project include the utilization of in *silico* docking to find novel targets to inhibit the

ribonuclease activity of the PIN domain during bacterial infections. Once targets are determined, these can be characterized in the *E. coli* growth recovery assay and the EpiAirways (Fig. 11 and 13) before advancing to the chinchilla model of experimental otitis media.

The second study designed a growth recovery assay similar to the *E. coli* model (Fig. 11) in the background of NTHi strain R2866 (Fig. 18). The toxin was cloned into a targeted delivery vector with the *araC* regulatory gene and was fused to the P_{BAD} promoter to study the induction of *vapD in cis* in the background of NTHi. *E.* coli arabinose permease genes were targeted into the chromosome of NTHi to determine if their presence increased the transport of arabinose across the membrane (Fig. 18). In the subsequent growth recovery assay, NTHi did not transport arabinose even with the help of on of two *E. coli* arabinose permeases tested *in cis*. The assay used in the present study was a good demonstration of the molecular techniques that can be utilized to study gene expression in NTHi. This study also serves as a reminder that one of the difficulties of vaccine construction for NTHi is the genetic heterogeneity among the strains. To successfully complete the assay design, multiple strains had to be used due to compatibility issues. In future studies, next steps include determining if lack of induction was due to insufficient amounts of arabinose transport across the membrane. To measure that, a permease could be cloned into an IPTG-inducible plasmid compatible in the background of R2866, as it has previously been determined that IPTG can cross the membrane of NTHi (146).

The third study investigated the prevalence and conservation of type II TA modules in 12 recently-sequenced NTHi. All 12 strains maintained at least three of the

four *vap* genes that were investigated. Type II TA loci contribute significantly to the survival and virulence of NTHi during infections and these loci increase bacterial persistence and dormancy following stress. Due to their conservation among strains and being essential for bacterial survival during infections, these TA loci are an important potential therapeutic target. Unlike other virulence factors such as antibiotic resistance genes and protective outer membrane proteins, which can mutate to adapt to antimicrobial therapy, it would be deleterious for bacteria to mutate these TA loci. Doing so would ultimately hasten cell death under stressful conditions.

In summary, the lack of newly-identified effective antibiotics increases the need to explore other non-standard molecular approaches to treat bacterial infections. Due to the conservation of the type II TA loci, these modules prove to be promising candidates for the development of a small molecule antimicrobial therapy.

REFERENCES

1. Pfeiffer R. 1892. Preliminary communication on the exciting causes of influenza. Br Med J 1:128.
2. Winslow CE, Broadhurst J, Buchanan RE, Krumwiede C, Rogers LA, Smith GH. 1917. The families and genera of the bacteria: preliminary report of the committee of the Society of American Bacteriologists on characterization and classification of bacterial types. J Bacteriol 2:505-566.
3. Pittman M. 1931. Variation and type specificity in the bacterial species *Haemophilus influenzae*. J Exp Med 53:471-492.
4. Kroll JS, Zamze S, Loynds B, Moxon ER. 1989. Common organization of chromosomal loci for production of different capsular polysaccharides in *Haemophilus influenzae*. J Bacteriol 171:3343-3347.
5. Kroll JS, Loynds BM, Moxon ER. 1991. The *Haemophilus influenzae* capsulation gene cluster: a compound transposon. Mol Microbiol 5:1549-1560.
6. Giufrè M, Cardines R, Mastrantonio P, Cerquetti M. 2010. Genetic characterization of the capsulation locus of *Haemophilus influenzae* serotype e. J Clin Microbiol 48:1404-1407.
7. Satola SW, Schirmer PL, Farley MM. 2003. Genetic analysis of the capsule locus of *Haemophilus influenzae* serotype f. Infect Immun 71:7202-7207.
8. Kroll JS, Loynds B, Brophy LN, Moxon ER. 1990. The *bex* locus in encapsulated *Haemophilus influenzae*: a chromosomal region involved in capsule polysaccharide export. Mol Microbiol 4:1853-1862.
9. Van Eldere J, Brophy L, Loynds B, Celis P, Hancock I, Carman S, Kroll JS, Moxon ER. 1995. Region II of the *Haemophilus influenzae* type b capsulation locus is involved in serotype-specific polysaccharide synthesis. Mol Microbiol 15:107-118.
10. Sukupolvi-Petty S, Grass S, St Geme JW. 2006. The *Haemophilus influenzae* type b *hcsA* and *hcsB* gene products facilitate transport of capsular polysaccharide across the outer membrane and are essential for virulence. J Bacteriol 188:3870-3877.
11. Murphy TV, Granoff DM, Pierson LM, Pastor P, White KE, Clements JF, Osterholm MT. 1992. Invasive *Haemophilus influenzae* type b disease in children less than 5 years of age in Minnesota and in Dallas County, Texas, 1983-1984. J Infect Dis 165 Suppl 1:S7-S10.
12. Adams WG, Deaver KA, Cochi SL, Plikaytis BD, Zell ER, Broome CV, Wenger JD. 1993. Decline of childhood *Haemophilus influenzae* type b (Hib) disease in the Hib vaccine era. JAMA 269:221-226.
13. Ladhani SN. 2012. Two decades of experience with the *Haemophilus influenzae* serotype b conjugate vaccine in the United Kingdom. Clin Ther 34:385-399.
14. Watt JP, Wolfson LJ, O'Brien KL, Henkle E, Deloria-Knoll M, McCall N, Lee E, Levine OS, Hajjeh R, Mulholland K, Cherian T, Team HaPGBoDS. 2009. Burden of disease caused by *Haemophilus influenzae* type b in children younger than 5 years: global estimates. Lancet 374:903-911.
15. Sriram KB, Cox AJ, Clancy RL, Slack MPE, Cripps AW. 2018. Nontypeable *Haemophilus influenzae* and chronic obstructive pulmonary disease: a review for clinicians. Crit Rev Microbiol 44:125-142.

16. Decramer M, Janssens W, Miravitlles M. 2012. Chronic obstructive pulmonary disease. Lancet 379:1341-1351.
17. Casey JR, Pichichero ME. 2004. Changes in frequency and pathogens causing acute otitis media in 1995-2003. Pediatr Infect Dis J 23:824-828.
18. Alsarraf R, Jung CJ, Perkins J, Crowley C, Alsarraf NW, Gates GA. 1999. Measuring the indirect and direct costs of acute otitis media. Arch Otolaryngol Head Neck Surg 125:12-18.
19. O'Brien MA, Prosser LA, Paradise JL, Ray GT, Kulldorff M, Kurs-Lasky M, Hinrichsen VL, Mehta J, Colborn DK, Lieu TA. 2009. New vaccines against otitis media: projected benefits and cost-effectiveness. Pediatrics 123:1452-1463.
20. Ahmed S, Shapiro NL, Bhattacharyya N. 2014. Incremental health care utilization and costs for acute otitis media in children. Laryngoscope 124:301-305.
21. Farjo RS, Foxman B, Patel MJ, Zhang L, Pettigrew MM, McCoy SI, Marrs CF, Gilsdorf JR. 2004. Diversity and sharing of *Haemophilus influenzae* strains colonizing healthy children attending day-care centers. Pediatr Infect Dis J 23:41-46.
22. Pettigrew MM, Ahearn CP, Gent JF, Kong Y, Gallo MC, Munro JB, D'Mello A, Sethi S, Tettelin H, Murphy TF. 2018. *Haemophilus influenzae* genome evolution during persistence in the human airways in chronic obstructive pulmonary disease. Proc Natl Acad Sci USA 115:E3256-E3265.
23. Cerquetti M, Giufrè M. 2016. Why we need a vaccine for non-typeable *Haemophilus influenzae*. Hum Vaccin Immunother 12:2357-2361.
24. Behrouzi A, Vaziri, Rahimi-Jamnani F, Afrough, Rahbar, Satarian, Siadat SD. 2017. Vaccine Candidates against nontypeable *Haemophilus influenzae*: a review. Iran Biomed J 21:69-76.
25. Apicella MA, Coffin J, Ketterer M, Post DMB, Day CJ, Jen FE, Jennings MP. 2018. Nontypeable *Haemophilus influenzae* lipooligosaccharide expresses a terminal ketodeoxyoctanoate. MBio 9.
26. Prymula R, Peeters P, Chrobok V, Kriz P, Novakova E, Kaliskova E, Kohl I, Lommel P, Poolman J, Prieels JP, Schuerman L. 2006. Pneumococcal capsular polysaccharides conjugated to protein D for prevention of acute otitis media caused by both *Streptococcus pneumoniae* and non-typable *Haemophilus influenzae*: a randomised double-blind efficacy study. Lancet 367:740-748.
27. Sveinsdóttir H, Björnsdóttir JB, Erlendsdóttir H, Hjálmarsdóttir M, Hrafnkelsson B, Haraldsson Á, Kristinsson KG, Haraldsson G. 2019. The effect of the 10-Valent pneumococcal nontypeable *Haemophilus influenzae* protein D conjugate vaccine on *H. influenzae* in healthy carriers and middle ear infections in Iceland. J Clin Microbiol 57.
28. van den Bergh MR, Spijkerman J, Swinnen KM, François NA, Pascal TG, Borys D, Schuerman L, Ijzerman EP, Bruin JP, van der Ende A, Veenhoven RH, Sanders EA. 2013. Effects of the 10-valent pneumococcal nontypeable *Haemophilus influenzae* protein D-conjugate vaccine on nasopharyngeal bacterial colonization in young children: a randomized controlled trial. Clin Infect Dis 56:e30-39.
29. Brandileone MC, Zanella RC, Almeida SCG, Brandao AP, Ribeiro AF, Carvalhanas TMP, Sato H, Andrade AL, Verani JR, Group PCS. 2016. Effect of

10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and *Haemophilus influenzae* among children in São Paulo, Brazil. Vaccine 34:5604-5611.

30. Gerdes K, Rasmussen PB, Molin S. 1986. Unique type of plasmid maintenance function: postsegregational killing of plasmid-free cells. Proc Natl Acad Sci USA 83:3116-3120.
31. Ogura T, Hiraga S. 1983. Mini-F plasmid genes that couple host cell division to plasmid proliferation. Proc Natl Acad Sci USA 80:4784-4788.
32. Masuda Y, Miyakawa K, Nishimura Y, Ohtsubo E. 1993. *chpA* and *chpB*, *Escherichia coli* chromosomal homologs of the *pem* locus responsible for stable maintenance of plasmid R100. J Bacteriol 175:6850-6856.
33. Gerdes K. 2000. Toxin-antitoxin modules may regulate synthesis of macromolecules during nutritional stress. J Bacteriol 182:561-572.
34. Yamaguchi Y, Park JH, Inouye M. 2011. Toxin-antitoxin systems in bacteria and archaea. Annu Rev Genet 45:61-79.
35. Correia FF, D'Onofrio A, Rejtar T, Li L, Karger BL, Makarova K, Koonin EV, Lewis K. 2006. Kinase activity of overexpressed HipA is required for growth arrest and multidrug tolerance in *Escherichia coli*. J Bacteriol 188:8360-8367.
36. Ren D, Walker AN, Daines DA. 2012. Toxin-antitoxin loci *vapBC-1* and *vapXD* contribute to survival and virulence in nontypeable *Haemophilus influenzae*. BMC Microbiol 12:263.
37. Ramage HR, Connolly LE, Cox JS. 2009. Comprehensive functional analysis of *Mycobacterium tuberculosis* toxin-antitoxin systems: implications for pathogenesis, stress responses, and evolution. PLoS Genet 5:e1000767.
38. Weaver KE, Reddy SG, Brinkman CL, Patel S, Bayles KW, Endres JL. 2009. Identification and characterization of a family of toxin-antitoxin systems related to the *Enterococcus faecalis* plasmid pAD1 par addiction module. Microbiology 155:2930-2940.
39. Christensen SK, Mikkelsen M, Pedersen K, Gerdes K. 2001. RelE, a global inhibitor of translation, is activated during nutritional stress. Proc Natl Acad Sci USA 98:14328-14333.
40. Tachdjian S, Kelly RM. 2006. Dynamic metabolic adjustments and genome plasticity are implicated in the heat shock response of the extremely thermoacidophilic archaeon *Sulfolobus solfataricus*. J Bacteriol 188:4553-4559.
41. Lemos JA, Brown TA, Abranches J, Burne RA. 2005. Characteristics of *Streptococcus mutans* strains lacking the *mazEF* and *relBE* toxin-antitoxin modules. FEMS Microbiol Lett 253:251-257.
42. Fernández-García L, Blasco L, Lopez M, Bou G, García-Contreras R, Wood T, Tomas M. 2016. Toxin-antitoxin systems in clinical pathogens. Toxins (Basel) 8.
43. Van Melderen L, Thi MH, Lecchi P, Gottesman S, Couturier M, Maurizi MR. 1996. ATP-dependent degradation of CcdA by Lon protease. Effects of secondary structure and heterologous subunit interactions. J Biol Chem 271:27730-27738.
44. Aizenman E, Engelberg-Kulka H, Glaser G. 1996. An *Escherichia coli* chromosomal "addiction module" regulated by guanosine 3',5'-bispyrophosphate:

a model for programmed bacterial cell death. Proc Natl Acad Sci USA 93:6059-6063.
45. Vogel J, Argaman L, Wagner EG, Altuvia S. 2004. The small RNA IstR inhibits synthesis of an SOS-induced toxic peptide. Curr Biol 14:2271-2276.
46. Daines DA, Wu MH, Yuan SY. 2007. VapC-1 of nontypeable *Haemophilus influenzae* is a ribonuclease. J Bacteriol 189:5041-5048.
47. Simic M, De Jonge N, Loris R, Vesnaver G, Lah J. 2009. Driving forces of gyrase recognition by the addiction toxin CcdB. J Biol Chem 284:20002-20010.
48. Mutschler H, Gebhardt M, Shoeman RL, Meinhart A. 2011. A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. PLoS Biol 9:e1001033.
49. Masuda H, Tan Q, Awano N, Wu KP, Inouye M. 2012. YeeU enhances the bundling of cytoskeletal polymers of MreB and FtsZ, antagonizing the CbtA (YeeV) toxicity in *Escherichia coli*. Mol Microbiol 84:979-989.
50. Fozo EM, Hemm MR, Storz G. 2008. Small toxic proteins and the antisense RNAs that repress them. Microbiol Mol Biol Rev 72:579-589.
51. Gerdes K, Thisted T, Martinussen J. 1990. Mechanism of post-segregational killing by the *hok/sok* system of plasmid R1: *sok* antisense RNA regulates formation of a *hok* mRNA species correlated with killing of plasmid-free cells. Mol Microbiol 4:1807-1818.
52. Fineran PC, Blower TR, Foulds IJ, Humphreys DP, Lilley KS, Salmond GP. 2009. The phage abortive infection system, *ToxIN*, functions as a protein-RNA toxin-antitoxin pair. Proc Natl Acad Sci USA 106:894-899.
53. Blower TR, Pei XY, Short FL, Fineran PC, Humphreys DP, Luisi BF, Salmond GP. 2011. A processed noncoding RNA regulates an altruistic bacterial antiviral system. Nat Struct Mol Biol 18:185-190.
54. Leplae R, Geeraerts D, Hallez R, Guglielmini J, Drèze P, Van Melderen L. 2011. Diversity of bacterial type II toxin-antitoxin systems: a comprehensive search and functional analysis of novel families. Nucleic Acids Res 39:5513-5525.
55. Winther KS, Gerdes K. 2011. Enteric virulence associated protein VapC inhibits translation by cleavage of initiator tRNA. Proc Natl Acad Sci USA 108:7403-7407.
56. Tan Q, Awano N, Inouye M. 2011. YeeV is an *Escherichia coli* toxin that inhibits cell division by targeting the cytoskeleton proteins, FtsZ and MreB. Mol Microbiol 79:109-118.
57. Wang X, Lord DM, Cheng HY, Osbourne DO, Hong SH, Sanchez-Torres V, Quiroga C, Zheng K, Herrmann T, Peti W, Benedik MJ, Page R, Wood TK. 2012. A new type V toxin-antitoxin system where mRNA for toxin GhoT is cleaved by antitoxin GhoS. Nat Chem Biol 8:855-861.
58. Aakre CD, Phung TN, Huang D, Laub MT. 2013. A bacterial toxin inhibits DNA replication elongation through a direct interaction with the β sliding clamp. Mol Cell 52:617-628.
59. Pandey DP, Gerdes K. 2005. Toxin-antitoxin loci are highly abundant in free-living but lost from host-associated prokaryotes. Nucleic Acids Res 33:966-976.

60. Ren D, Kordis AA, Sonenshine DE, Daines DA. 2014. The *toxAvapA* toxin-antitoxin locus contributes to the survival of nontypeable *Haemophilus influenzae* during infection. PLoS One 9:e91523.
61. Wall D, Kaiser D. 1999. Type IV pili and cell motility. Mol Microbiol 32:1-10.
62. Bendtsen KL, Brodersen DE. 2017. Higher-order structure in bacterial *vapBC* toxin-antitoxin complexes. Subcell Biochem 83:381-412.
63. Liu M, Zhang Y, Inouye M, Woychik NA. 2008. Bacterial addiction module toxin Doc inhibits translation elongation through its association with the 30S ribosomal subunit. Proc Natl Acad Sci USA 105:5885-5890.
64. Masuda H, Inouye M. 2017. Toxins of prokaryotic toxin-antitoxin systems with sequence-specific endoribonuclease activity. Toxins (Basel) 9.
65. Arcus VL, McKenzie JL, Robson J, Cook GM. 2011. The PIN-domain ribonucleases and the prokaryotic VapBC toxin-antitoxin array. Protein Eng Des Sel 24:33-40.
66. Coussens NP, Daines DA. 2016. Wake me when it's over - bacterial toxin-antitoxin proteins and induced dormancy. Exp Biol Med (Maywood) 241:1332-1342.
67. Deep A, Kaundal S, Agarwal S, Singh R, Thakur KG. 2017. Crystal structure of *Mycobacterium tuberculosis* VapC20 toxin and its interactions with cognate antitoxin, VapB20, suggest a model for toxin-antitoxin assembly. FEBS J 284:4066-4082.
68. Krissinel E, Henrick K. 2007. Inference of macromolecular assemblies from crystalline state. J Mol Biol 372:774-797.
69. Krissinel E, Henrick K. 2004. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 60:2256-2268.
70. Xu K, Dedic E, Brodersen DE. 2016. Structural analysis of the active site architecture of the VapC toxin from *Shigella flexneri*. Proteins 84:892-899.
71. Dienemann C, Bøggild A, Winther KS, Gerdes K, Brodersen DE. 2011. Crystal structure of the VapBC toxin-antitoxin complex from *Shigella flexneri* reveals a hetero-octameric DNA-binding assembly. J Mol Biol 414:713-722.
72. Maté MJ, Vincentelli R, Foos N, Raoult D, Cambillau C, Ortiz-Lombardía M. 2012. Crystal structure of the DNA-bound VapBC2 antitoxin/toxin pair from *Rickettsia felis*. Nucleic Acids Res 40:3245-3258.
73. Senissar M, Manav MC, Brodersen DE. 2017. Structural conservation of the PIN domain active site across all domains of life. Protein Sci 26:1474-1492.
74. Daines DA, Silver RP. 2000. Evidence for multimerization of neu proteins involved in polysialic acid synthesis in *Escherichia coli* K1 using improved LexA-based vectors. J Bacteriol 182:5267-5270.
75. Daines DA, Granger-Schnarr M, Dimitrova M, Silver RP. 2002. Use of LexA-based system to identify protein-protein interactions in vivo. Methods Enzymol 358:153-161.
76. Miller JH. 1972. Experiments in Molecular Genetics. Cold Spring Harbor Laboratory, NY:352-355.

77. Guzman LM, Belin D, Carson MJ, Beckwith J. 1995. Tight regulation, modulation, and high-level expression by vectors containing the arabinose P_{BAD} promoter. J Bacteriol 177:4121-4130.
78. Brosius J, Erfle M, Storella J. 1985. Spacing of the -10 and -35 regions in the *tac* promoter. Effect on its *in vivo* activity. J Biol Chem 260:3539-3541.
79. Klein RD, Luginbuhl GH. 1979. Simplified media for the growth of *Haemophilus influenzae* from clinical and normal flora sources. J Gen Microbiol 113:409-411.
80. Herriott RM, Meyer EY, Vogt M, Modan M. 1970. Defined medium for growth of *Haemophilus influenzae*. J Bacteriol 101:513-516.
81. Herriott RM, Meyer EM, Vogt M. 1970. Defined nongrowth media for stage II development of competence in *Haemophilus influenzae*. J Bacteriol 101:517-524.
82. Hamilton B, Manzella A, Schmidt K, DiMarco V, Butler JS. 2014. Analysis of non-typeable *Haemophilus influenzae* VapC-1 mutations reveals structural features required for toxicity and flexibility in the active site. PLoS One 9:e112921.
83. Zaychikova MV, Mikheecheva NE, Belay YO, Alekseeva MG, Melerzanov AV, Danilenko VN. 2018. Single nucleotide polymorphisms of Beijing lineage *Mycobacterium tuberculosis* toxin-antitoxin system genes: their role in the changes of protein activity and evolution. Tuberculosis (Edinb) 112:11-19.
84. Cline SD, Saleem S, Daines DA. 2012. Regulation of the *vapBC-1* toxin-antitoxin locus in nontypeable *Haemophilus influenzae*. PLoS One 7:e32199.
85. Daines DA, Jarisch J, Smith AL. 2004. Identification and characterization of a nontypeable *Haemophilus influenzae* putative toxin-antitoxin locus. BMC Microbiol 4:30.
86. Agarwal S, Tiwari P, Deep A, Kidwai S, Gupta S, Thakur KG, Singh R. 2018. System-wide analysis unravels the differential regulation and *in vivo* essentiality of virulence-associated proteins B and C toxin-antitoxin systems of *Mycobacterium tuberculosis*. J Infect Dis 217:1809-1820.
87. Wen W, Liu B, Xue L, Zhu Z, Niu L, Sun B. 2018. Autoregulation and virulence control by the toxin-antitoxin system *savRS* in *Staphylococcus aureus*. Infect Immun 86.
88. Gotfredsen M, Gerdes K. 1998. The *Escherichia coli relBE* genes belong to a new toxin-antitoxin gene family. Mol Microbiol 29:1065-1076.
89. Deep A, Tiwari P, Agarwal S, Kaundal S, Kidwai S, Singh R, Thakur KG. 2018. Structural, functional and biological insights into the role of *Mycobacterium tuberculosis vapBC11* toxin-antitoxin system: targeting a tRNase to tackle mycobacterial adaptation. Nucleic Acids Res 46:11639-11655.
90. Kabsch W. 1988. Automatic indexing of rotation diffraction patterns. Journal of Applied Crystallography 21:67-72.
91. Kabsch W. 2010. XDS. Acta Crystallogr D Biol Crystallogr 66:125-132.
92. Vonrhein C, Flensburg C, Keller P, Sharff A, Smart O, Paciorek W, Womack T, Bricogne G. 2011. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr D Biol Crystallogr 67:293-302.
93. Evans PR. 2011. An introduction to data reduction: space-group determination, scaling and intensity statistics. Acta Crystallogr D Biol Crystallogr 67:282-292.
94. Vagin A, Lebedev A. 2015. MoRDa, an automatic molecular replacement pipeline. Acta Crystallographica Section A 71:s19.

95. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. *Phaser* crystallographic software. J Appl Cryst 40:658-674.
96. Langer G, Cohen SX, Lamzin VS, Perrakis A. 2008. Automated macromolecular model building for x-ray crystallography using ARP/wARP version 7. Nat Protoc 3:1171-1179.
97. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66:213-221.
98. Emsley P, Lohkamp B, Scott WG, Cowtan K. 2010. Features and development of Coot. Acta Crystallogr D Biol Crystallogr 66:486-501.
99. Chen VB, Arendall WB, 3rd, Headd JJ, Keedy DA, Immormino RM, Kapral GJ, Murray LW, Richardson JS, Richardson DC. 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr D Biol Crystallogr 66:12-21.
100. Potterton L, McNicholas S, Krissinel E, Gruber J, Cowtan K, Emsley P, Murshudov GN, Cohen S, Perrakis A, Noble M. 2004. Developments in the CCP4 molecular-graphics project. Acta Crystallogr D Biol Crystallogr 60:2288-2294.
101. Krissinel E. 2012. Enhanced fold recognition using efficient short fragment clustering. Journal of Molecular Biochemistry 1:76-85.
102. Evans P. 2006. Scaling and assessment of data quality. Acta Crystallogr D Biol Crystallogr 62:72-82.
103. Diederichs K, Karplus PA. 1997. Improved R-factors for diffraction data analysis in macromolecular crystallography. Nat Struct Biol 4:269-275.
104. Karplus PA, Diederichs K. 2012. Linking crystallographic model and data quality. Science 336:1030-1033.
105. Evans P. 2012. Biochemistry. Resolving some old problems in protein crystallography. Science 336:986-987.
106. Daines DA, Smith AL. 2004. Construction of a nontypeable *Haemophilus influenzae*-specific ectopic delivery vector. J Microbiol Methods 57:421-424.
107. Robinson E, Juhas M, Hood D, Crook D. 2010. Construction of a novel shuttle vector for use in *Haemophilus influenzae* and *H. parainfluenzae*. J Microbiol Methods 83:330-334.
108. Ren D, Daines DA. 2011. Use of the EpiAirway model for characterizing long-term host-pathogen interactions. J Vis Exp:e3261.
109. Backman K, Ptashne M, Gilbert W. 1976. Construction of plasmids carrying the *cI* gene of bacteriophage lambda. Proc Natl Acad Sci USA 73:4174-4178.
110. Roberts TM, Kacich R, Ptashne M. 1979. A general method for maximizing the expression of a cloned gene. Proc Natl Acad Sci USA 76:760-764.
111. Yanisch-Perron C, Vieira J, Messing J. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33:103-119.
112. de Boer HA, Comstock LJ, Vasser M. 1983. The *tac* promoter: a functional hybrid derived from the *trp* and *lac* promoters. Proc Natl Acad Sci USA 80:21-25.

113. Amann E, Brosius J, Ptashne M. 1983. Vectors bearing a hybrid *trp-lac* promoter useful for regulated expression of cloned genes in *Escherichia coli*. Gene 25:167-178.
114. Tabor S. 2001. Expression using the T7 RNA polymerase/promoter system. Curr Protoc Mol Biol Chapter 16:Unit16.2.
115. Stoner C, Schleif R. 1983. The *araE* low affinity L-arabinose transport promoter. Cloning, sequence, transcription start site and DNA binding sites of regulatory proteins. J Mol Biol 171:369-381.
116. Daruwalla KR, Paxton AT, Henderson PJ. 1981. Energization of the transport systems for arabinose and comparison with galactose transport in *Escherichia coli*. Biochem J 200:611-627.
117. Schleif R. 1969. An L-arabinose binding protein and arabinose permeation in *Escherichia coli*. J Mol Biol 46:185-196.
118. Brown CE, Hogg RW. 1972. A second transport system for L-arabinose in *Escherichia coli* B-r controlled by the *araC* gene. J Bacteriol 111:606-613.
119. Hogg RW, Englesberg E. 1969. L-arabinose binding protein from *Escherichia coli* B-r. J Bacteriol 100:423-432.
120. Horazdovsky BF, Hogg RW. 1987. High-affinity L-arabinose transport operon. Gene product expression and mRNAs. J Mol Biol 197:27-35.
121. Schleif R. 2000. Regulation of the L-arabinose operon of *Escherichia coli*. Trends Genet 16:559-565.
122. Helling RB, Weinberg R. 1963. Complementation studies of arabinose genes in *Escherichia coli*. Genetics 48:1397-1410.
123. Greenblatt J, Schleif R. 1971. Arabinose C protein: regulation of the arabinose operon *in vitro*. Nat New Biol 233:166-170.
124. Johnson CM, Schleif RF. 1995. *In vivo* induction kinetics of the arabinose promoters in *Escherichia coli*. J Bacteriol 177:3438-3442.
125. Casadaban MJ. 1976. Regulation of the regulatory gene for the arabinose pathway, *araC*. J Mol Biol 104:557-566.
126. Dunn TM, Hahn S, Ogden S, Schleif RF. 1984. An operator at -280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc Natl Acad Sci USA 81:5017-5020.
127. Martin K, Huo L, Schleif RF. 1986. The DNA loop model for *ara* repression: AraC protein occupies the proposed loop sites *in vivo* and repression-negative mutations lie in these same sites. Proc Natl Acad Sci USA 83:3654-3658.
128. Soisson SM, MacDougall-Shackleton B, Schleif R, Wolberger C. 1997. Structural basis for ligand-regulated oligomerization of AraC. Science 276:421-425.
129. Saviola B, Seabold R, Schleif RF. 1998. Arm-domain interactions in AraC. J Mol Biol 278:539-548.
130. Schleif R. 1992. DNA looping. Annu Rev Biochem 61:199-223.
131. Schleif R. 2010. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. FEMS Microbiol Rev 34:779-796.
132. Schleif R. 2003. AraC protein: a love-hate relationship. Bioessays 25:274-282.

133. Lobell RB, Schleif RF. 1990. DNA looping and unlooping by AraC protein. Science 250:528-532.
134. Harmer T, Wu M, Schleif R. 2001. The role of rigidity in DNA looping-unlooping by AraC. Proc Natl Acad Sci USA 98:427-431.
135. Dunn TM, Schleif R. 1984. Deletion analysis of the *Escherichia coli ara* P_C and P_{BAD} promoters. J Mol Biol 180:201-204.
136. Fang FC, Chen CY, Guiney DG, Xu Y. 1996. Identification of sigma S-regulated genes in *Salmonella typhimurium*: complementary regulatory interactions between sigma S and cyclic AMP receptor protein. J Bacteriol 178:5112-5120.
137. Newman JR, Fuqua C. 1999. Broad-host-range expression vectors that carry the L-arabinose-inducible *Escherichia coli araBAD* promoter and the *araC* regulator. Gene 227:197-203.
138. Ben-Samoun K, Leblon G, Reyes O. 1999. Positively regulated expression of the *Escherichia coli araBAD* promoter in *Corynebacterium glutamicum*. FEMS Microbiol Lett 174:125-130.
139. Sukchawalit R, Vattanaviboon P, Sallabhan R, Mongkolsuk S. 1999. Construction and characterization of regulated L-arabinose-inducible broad host range expression vectors in *Xanthomonas*. FEMS Microbiol Lett 181:217-223.
140. Damron FH, McKenney ES, Schweizer HP, Goldberg JB. 2013. Construction of a broad-host-range Tn7-based vector for single-copy P_{BAD}-controlled gene expression in Gram-negative bacteria. Appl Environ Microbiol 79:718-721.
141. Meibom KL, Blokesch M, Dolganov NA, Wu CY, Schoolnik GK. 2005. Chitin induces natural competence in *Vibrio cholerae*. Science 310:1824-1827.
142. Lefebre MD, Valvano MA. 2002. Construction and evaluation of plasmid vectors optimized for constitutive and regulated gene expression in *Burkholderia cepacia* complex isolates. Appl Environ Microbiol 68:5956-5964.
143. Molinaro AL, Kashipathy MM, Lovell S, Battaile KP, Coussens NP, Shen M, Daines DA. 2019. Crystal structure of VapBC-1 from nontypeable *Haemophilus influenzae* and the effect of PIN domain mutations on survival during infection. J Bacteriol 201.
144. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389-3402.
145. Coussens NP, Molinaro AL, Culbertson KJ, Peryea T, Zahoránszky-Köhalmi G, Hall MD, Daines DA. 2018. Better living through chemistry: addressing emerging antibiotic resistance. Exp Biol Med (Maywood):1535370218755659.
146. Daines DA, Smith AL. 2001. Design and construction of a *Haemophilus influenzae* conjugal expression system. Gene 281:95-102.
147. Coleman HN, Daines DA, Jarisch J, Smith AL. 2003. Chemically defined media for growth of *Haemophilus influenzae* strains. J Clin Microbiol 41:4408-4410.
148. Crameri A, Whitehorn EA, Tate E, Stemmer WP. 1996. Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat Biotechnol 14:315-319.
149. Wong SM, Akerley BJ. 2003. Inducible expression system and marker-linked mutagenesis approach for functional genomics of *Haemophilus influenzae*. Gene 316:177-186.

150. Shimoyama M, Smith JR, De Pons J, Tutaj M, Khampang P, Hong W, Erbe CB, Ehrlich GD, Bakaletz LO, Kerschner JE. 2016. The Chinchilla Research Resource Database: resource for an otolaryngology disease model. Database (Oxford) 2016.
151. Moleres J, Fernández-Calvet A, Ehrlich RL, Martí S, Pérez-Regidor L, Euba B, Rodríguez-Arce I, Balashov S, Cuevas E, Liñares J, Ardanuy C, Martín-Santamaría S, Ehrlich GD, Mell JC, Garmendia J. 2018. Antagonistic pleiotropy in the bifunctional surface protein FadL (OmpP1) during adaptation of *Haemophilus influenzae* to chronic lung infection associated with chronic obstructive pulmonary disease. MBio 9.
152. Macfadyen LP, Dorocicz IR, Reizer J, Saier MH, Redfield RJ. 1996. Regulation of competence development and sugar utilization in *Haemophilus influenzae* Rd by a phosphoenolpyruvate:fructose phosphotransferase system. Mol Microbiol 21:941-952.
153. Mell JC, Lee JY, Firme M, Sinha S, Redfield RJ. 2014. Extensive cotransformation of natural variation into chromosomes of naturally competent *Haemophilus influenzae*. G3 (Bethesda) 4:717-731.
154. Long SS, Henretig FM, Teter MJ, McGowan KL. 1983. Nasopharyngeal flora and acute otitis media. Infect Immun 41:987-991.
155. Langereis JD, de Jonge MI. 2015. Invasive disease caused by nontypeable *Haemophilus influenzae*. Emerg Infect Dis 21:1711-1718.
156. Muda NM, Nasreen M, Dhouib R, Hosmer J, Hill J, Mahawar M, Schirra HJ, McEwan AG, Kappler U. 2019. Metabolic analyses reveal common adaptations in two invasive *Haemophilus influenzae* strains. Pathog Dis 77.
157. Erwin AL, Smith AL. 2007. Nontypeable *Haemophilus influenzae*: understanding virulence and commensal behavior. Trends Microbiol 15:355-362.
158. King P. 2012. *Haemophilus influenzae* and the lung (*Haemophilus* and the lung). Clin Transl Med 1:10.
159. Sakka L, Coll G, Chazal J. 2011. Anatomy and physiology of cerebrospinal fluid. Eur Ann Otorhinolaryngol Head Neck Dis 128:309-316.
160. Wong SM, Akerley BJ. 2012. Genome-scale approaches to identify genes essential for *Haemophilus influenzae* pathogenesis. Front Cell Infect Microbiol 2:23.
161. Othman DS, Schirra H, McEwan AG, Kappler U. 2014. Metabolic versatility in *Haemophilus influenzae*: a metabolomic and genomic analysis. Front Microbiol 5:69.
162. Dhouib R, Othman DS, Lin V, Lai XJ, Wijesinghe HG, Essilfie AT, Davis A, Nasreen M, Bernhardt PV, Hansbro PM, McEwan AG, Kappler U. 2016. A novel, molybdenum-containing methionine sulfoxide reductase supports survival of *Haemophilus influenzae* in an *in vivo* model of infection. Front Microbiol 7:1743.
163. Macfadyen LP, Redfield RJ. 1996. Life in mucus: sugar metabolism in *Haemophilus influenzae*. Res Microbiol 147:541-551.
164. Kok M, Bron G, Erni B, Mukhija S. 2003. Effect of enzyme I of the bacterial phosphoenolpyruvate : sugar phosphotransferase system (PTS) on virulence in a murine model. Microbiology 149:2645-2652.

165. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol 215:403-410.
166. Bakaletz LO, Baker BD, Jurcisek JA, Harrison A, Novotny LA, Bookwalter JE, Mungur R, Munson RS. 2005. Demonstration of type IV pilus expression and a twitching phenotype by *Haemophilus influenzae*. Infect Immun 73:1635-1643.
167. Elbing KL, Brent R. 2019. Growth of *E. coli* in liquid medium. Curr Protoc Mol Biol 125:e81.
168. Fleischmann RD, Adams MD, White O, Clayton RA, Kirkness EF, Kerlavage AR, Bult CJ, Tomb JF, Dougherty BA, Merrick JM. 1995. Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. Science 269:496-512.
169. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. 2017. GenBank. Nucleic Acids Res 45:D37-D42.
170. Besser J, Carleton HA, Gerner-Smidt P, Lindsey RL, Trees E. 2018. Next-generation sequencing technologies and their application to the study and control of bacterial infections. Clin Microbiol Infect 24:335-341.
171. Ma Z, Geng J, Yi L, Xu B, Jia R, Li Y, Meng Q, Fan H, Hu S. 2013. Insight into the specific virulence related genes and toxin-antitoxin virulent pathogenicity islands in swine streptococcosis pathogen *Streptococcus equi* ssp. zooepidemicus strain ATCC35246. BMC Genomics 14:377.
172. Schmidt H, Hensel M. 2004. Pathogenicity islands in bacterial pathogenesis. Clin Microbiol Rev 17:14-56.
173. Meats E, Feil EJ, Stringer S, Cody AJ, Goldstein R, Kroll JS, Popovic T, Spratt BG. 2003. Characterization of encapsulated and noncapsulated *Haemophilus influenzae* and determination of phylogenetic relationships by multilocus sequence typing. J Clin Microbiol 41:1623-1636.
174. Sacchi CT, Alber D, Dull P, Mothershed EA, Whitney AM, Barnett GA, Popovic T, Mayer LW. 2005. High level of sequence diversity in the 16S rRNA genes of *Haemophilus influenzae* isolates is useful for molecular subtyping. J Clin Microbiol 43:3734-3742.
175. Power PM, Bentley SD, Parkhill J, Moxon ER, Hood DW. 2012. Investigations into genome diversity of *Haemophilus influenzae* using whole genome sequencing of clinical isolates and laboratory transformants. BMC Microbiol 12:273.
176. Munson RS, Harrison A, Gillaspy A, Ray WC, Carson M, Armbruster D, Gipson J, Gipson M, Johnson L, Lewis L, Dyer DW, Bakaletz LO. 2004. Partial analysis of the genomes of two nontypeable *Haemophilus influenzae* otitis media isolates. Infect Immun 72:3002-3010.
177. Suzuki K, Bakaletz LO. 1994. Synergistic effect of adenovirus type 1 and nontypeable *Haemophilus influenzae* in a chinchilla model of experimental otitis media. Infect Immun 62:1710-1718.
178. Ren D, Nelson KL, Uchakin PN, Smith AL, Gu XX, Daines DA. 2012. Characterization of extended co-culture of non-typeable *Haemophilus influenzae* with primary human respiratory tissues. Exp Biol Med (Maywood) 237:540-547.
179. Boratyn GM, Camacho C, Cooper PS, Coulouris G, Fong A, Ma N, Madden TL, Matten WT, McGinnis SD, Merezhuk Y, Raytselis Y, Sayers EW, Tao T, Ye J,

Zaretskaya I. 2013. BLAST: a more efficient report with usability improvements. Nucleic Acids Res 41:W29-W33.
180. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, Lopez R, McWilliam H, Remmert M, Söding J, Thompson JD, Higgins DG. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol 7:539.
181. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R. 2010. A new bioinformatics analysis tools framework at EMBL-EBI. Nucleic Acids Res 38:W695-W699.
182. McWilliam H, Li W, Uludag M, Squizzato S, Park YM, Buso N, Cowley AP, Lopez R. 2013. Analysis tool web services from the EMBL-EBI. Nucleic Acids Res 41:W597-W600.
183. Tian QB, Hayashi T, Murata T, Terawaki Y. 1996. Gene product identification and promoter analysis of *hig* locus of plasmid Rts1. Biochem Biophys Res Commun 225:679-684.
184. Davis JJ, Gerdes S, Olsen GJ, Olson R, Pusch GD, Shukla M, Vonstein V, Wattam AR, Yoo H. 2016. PATtyFams: protein families for the microbial genomes in the PATRIC database. Front Microbiol 7:118.
185. Letunic I, Bork P. 2019. Interactive Tree Of Life (iTOL) v4: recent updates and new developments. Nucleic Acids Res.
186. Campagnari AA, Gupta MR, Dudas KC, Murphy TF, Apicella MA. 1987. Antigenic diversity of lipooligosaccharides of nontypable *Haemophilus influenzae*. Infect Immun 55:882-887.
187. Smith-Vaughan HC, Chang AB, Sarovich DS, Marsh RL, Grimwood K, Leach AJ, Morris PS, Price EP. 2014. Absence of an important vaccine and diagnostic target in carriage and disease-related nontypeable *Haemophilus influenzae*. Clin Vaccine Immunol 21:250-252.
188. Shen K, Antalis P, Gladitz J, Sayeed S, Ahmed A, Yu S, Hayes J, Johnson S, Dice B, Dopico R, Keefe R, Janto B, Chong W, Goodwin J, Wadowsky RM, Erdos G, Post JC, Ehrlich GD, Hu FZ. 2005. Identification, distribution, and expression of novel genes in 10 clinical isolates of nontypeable *Haemophilus influenzae*. Infect Immun 73:3479-3491.
189. Iskander M, Hayden K, Van Domselaar G, Tsang R. 2017. First complete genome sequence of *Haemophilus influenzae* serotype a. Genome Announc 5.
190. Giufrè M, Cardines R, Brigante G, Orecchioni F, Cerquetti M. 2017. Emergence of invasive *Haemophilus influenzae* type a disease in Italy. Clin Infect Dis 64:1626-1628.
191. Whitby PW, Seale TW, Morton DJ, VanWagoner TM, Stull TL. 2010. Characterization of the *Haemophilus influenzae tehB* gene and its role in virulence. Microbiology 156:1188-1200.
192. Giufrè M, Cardines R, Cerquetti M. 2017. First whole-genome sequence of a *Haemophilus influenzae* type e strain isolated from a patient with invasive disease in Italy. Genome Announc 5.
193. Su YC, Hörhold F, Singh B, Riesbeck K. 2013. Complete genome sequence of encapsulated *Haemophilus influenzae* type f KR494, an invasive isolate that caused necrotizing myositis. Genome Announc 1.

194. Bateman AC, Perez-Osorio AC, Li Z, Tran M, Greninger AL. 2017. Conservation and recombination in the genome sequence of *Haemophilus influenzae* type f WAPHL1. Genome Announc 5.
195. Watts SC, Holt KE. 2019. hicap: *in silico* sertotyping of the *Haemophilus influenzae* capsule locus. J Clin Microbiol 57.
196. Tatusov RL, Mushegian AR, Bork P, Brown NP, Hayes WS, Borodovsky M, Rudd KE, Koonin EV. 1996. Metabolism and evolution of *Haemophilus influenzae* deduced from a whole-genome comparison with *Escherichia coli*. Curr Biol 6:279-291.
197. Nielsen SM, de Gier C, Dimopoulou C, Gupta V, Hansen LH, Nørskov-Lauritsen N. 2015. The capsule biosynthesis locus of *Haemophilus influenzae* shows conspicuous similarity to the corresponding locus in *Haemophilus sputorum* and may have been recruited from this species by horizontal gene transfer. Microbiology 161:1182-1188.
198. Connor TR, Corander J, Hanage WP. 2012. Population subdivision and the detection of recombination in non-typable *Haemophilus influenzae*. Microbiology 158:2958-2964.
199. Musser JM, Barenkamp SJ, Granoff DM, Selander RK. 1986. Genetic relationships of serologically nontypable and serotype b strains of *Haemophilus influenzae*. Infect Immun 52:183-191.
200. Garmendia J, Viadas C, Calatayud L, Mell JC, Martí-Lliteras P, Euba B, Llobet E, Gil C, Bengoechea JA, Redfield RJ, Liñares J. 2014. Characterization of nontypable *Haemophilus influenzae* isolates recovered from adult patients with underlying chronic lung disease reveals genotypic and phenotypic traits associated with persistent infection. PLoS One 9:e97020.
201. Sender R, Fuchs S, Milo R. 2016. Revised estimates for the number of human and bacteria cells in the body. PLoS Biol 14:e1002533.
202. Otto M. 2014. Physical stress and bacterial colonization. FEMS Microbiol Rev 38:1250-1270.
203. Sassone-Corsi M, Raffatellu M. 2015. No vacancy: how beneficial microbes cooperate with immunity to provide colonization resistance to pathogens. J Immunol 194:4081-4087.
204. Kaur R, Chang A, Xu Q, Casey JR, Pichichero ME. 2011. Phylogenetic relatedness and diversity of non-typable *Haemophilus influenzae* in the nasopharynx and middle ear fluid of children with acute otitis media. J Med Microbiol 60:1841-1848.
205. Kędzierska B, Hayes F. 2016. Emerging roles of toxin-antitoxin modules in bacterial pathogenesis. Molecules 21.
206. Harms A, Brodersen DE, Mitarai N, Gerdes K. 2018. Toxins, targets, and triggers: an overview of toxin-antitoxin biology. Mol Cell 70:768-784.
207. Georgiades K, Raoult D. 2011. Genomes of the most dangerous epidemic bacteria have a virulence repertoire characterized by fewer genes but more toxin-antitoxin modules. PLoS One 6:e17962.
208. Xie Y, Wei Y, Shen Y, Li X, Zhou H, Tai C, Deng Z, Ou HY. 2018. TADB 2.0: an updated database of bacterial type II toxin-antitoxin loci. Nucleic Acids Res 46:D749-D753.

209. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, McNeil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. 2008. The RAST server: rapid annotations using subsystems technology. BMC Genomics 9:75.
210. Overbeek R, Olson R, Pusch GD, Olsen GJ, Davis JJ, Disz T, Edwards RA, Gerdes S, Parrello B, Shukla M, Vonstein V, Wattam AR, Xia F, Stevens R. 2014. The SEED and the rapid annotation of microbial genomes using subsystems technology (RAST). Nucleic Acids Res 42:D206-D214.
211. Wattam AR, Davis JJ, Assaf R, Boisvert S, Brettin T, Bun C, Conrad N, Dietrich EM, Disz T, Gabbard JL, Gerdes S, Henry CS, Kenyon RW, Machi D, Mao C, Nordberg EK, Olsen GJ, Murphy-Olson DE, Olson R, Overbeek R, Parrello B, Pusch GD, Shukla M, Vonstein V, Warren A, Xia F, Yoo H, Stevens RL. 2017. Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res 45:D535-D542.
212. Edgar RC. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32:1792-1797.
213. Cock PJ, Antao T, Chang JT, Chapman BA, Cox CJ, Dalke A, Friedberg I, Hamelryck T, Kauff F, Wilczynski B, de Hoon MJ. 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25:1422-1423.
214. Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312-1313.
215. Stamatakis A, Hoover P, Rougemont J. 2008. A rapid bootstrap algorithm for the RAxML Web servers. Syst Biol 57:758-771.

COPYRIGHTS

Overview and Crystal Structure of VapBC-1 from Nontypeable *Haemophilus influenzae* and the Effect of PIN Domain Mutations on Survival during Infection

The content used in the overview and chapter 1 is published in *Journal of Bacteriology*. Text and figures were modified to fit according to chapter format and the Old Dominion University Guide for Preparation of Theses and Dissertations.

FIG 1

This figure is published in the Journal of Clinical Microbiology (part of American Society for Microbiology journals), Title: hicap*:* in silico *serotyping of the Haemophilus influenzae capsule locus* (https://jcm.asm.org/content/57/6/e00190-19). This article is distributed under the terms of the Creative Commons Attribution 4.0 License (http://www.creativecommons.org/licenses/by-nc/4.0/) which permits non-commercial

use, reproduction and distribution of the work without further permission provided the original work is attributed as specified.

FIG 2

This figure is published in PLOS ONE, Title: *Regulation of the vapBC-1 toxin-antitoxin locus in nontypeable Haemophilus influenzae* (https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0032199). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are properly credited.

FIG 14

License Details

This Agreement between Ashley Molinaro ("You") and John Wiley and Sons ("John Wiley and Sons") consists of your license details and the terms and conditions provided by John Wiley and Sons and Copyright Clearance Center.

Print Copy

License Number	4638311474252
License date	Jul 29, 2019
Licensed Content Publisher	John Wiley and Sons
Licensed Content Publication	BioEssays
Licensed Content Title	AraC protein: A love–hate relationship
Licensed Content Author	Robert Schleif
Licensed Content Date	Feb 20, 2003
Licensed Content Volume	25
Licensed Content Issue	3
Licensed Content Pages	9
Type of Use	Dissertation/Thesis
Requestor type	University/Academic
Format	Print and electronic
Portion	Figure/table
Number of figures/tables	1
Original Wiley figure/table number(s)	Figure 2
Will you be translating?	No
Title of your thesis / dissertation	TARGETING THE PIN DOMAIN OF TYPE II TOXINS IS A NOVEL APPROACH TO TREAT INFECTIONS CAUSED BY NONTYPEABLE HAEMOPHILUS INFLUENZAE
Expected completion date	Sep 2019
Expected size (number of pages)	133
Requestor Location	Ashley Molinaro 2480 BRUSH CREEK LN Virginia Beach, VA 23454 United States Attn:
Publisher Tax ID	EU826007151
Total	**0.00 USD**

FIG 21

VITA

ASHLEY LYNNE MOLINARO

Graduate School, 2102 Monarch Hall, Old Dominion University.

EDUCATION

- BS: Biology, Minor: Criminal Justice, Norwich University, Northfield, VT

PEER-REVIEWED PUBLICATIONS

- Molinaro AL, Kashipathy MM, Lovell S, Battaile KP, Coussens NP, Shen M, Daines DA. 2019. Crystal Structure of VapBC-1 from Nontypeable *Haemophilus influenzae* and the Effect of PIN Domain Mutations on Survival during Infection. J Bacteriol 201.
- Coussens NP, Molinaro AL, Culbertson KJ, Peryea T, Zahoránszky-Köhalmi G, Hall MD, Daines DA. 2018. Better living through chemistry: Addressing emerging antibiotic resistance. Exp Biol Med (Maywood):1535370218755659.

PRESENTATIONS AND WORKSHOPS

- PATRIC RAST Workshop at Argonne National Laboratory, (2018)
- "Investigating the Role of Tetracycline Derivatives in Gram-Negative Metabolic Regulation" American Society of Microbiology, Virginia Branch, Oral presentation, (2017)
- "Investigating an Alternate Mechanism of Action for Methacycline" American Society of Microbiology/European Society of Clinical Microbiology, Poster 74, (2017)
- Assay Guidance Workshop for High-Throughput Screen and Lead Discovery National Center for Advancing Translational Sciences, (2017)
- "Noncognate Toxin-Antitoxin Interactions: Implications for Bacterial Persistence Experimental Biology, Poster C335, (2016)

HONORS AND AWARDS

- Virginia S. Bagley Endowed Scholarship (2018/2019)
- Old Dominion University Graduate Student Research Travel Award (2018)
- 2nd Place Presentation Award, American Society for Microbiology, Virginia Branch Annual Meeting (2017)
- Old Dominion University Candidate for the 68th Lindau Nobel Laureate Meeting (2017)
- American Society of Microbiology/European Society of Clinical Microbiology and Infectious Diseases Travel Award (2017)
- Biology Graduate Student Organization Travel Award (2017)
- Old Dominion University Outstanding Student Leader (2015)